Document made available under the Patent Cooperation Treaty (PCT)

International application number: PCT/GB05/000752

International filing date: 01 March 2005 (01.03.2005)

Document type: Certified copy of priority document

Document details: Country/Office: GB
Number: 0404574.6
Filing date: 01 March 2004 (01.03.2004)

Date of receipt at the International Bureau: 07 April 2005 (07.04.2005)

Remark: Priority document submitted or transmitted to the International Bureau in compliance with Rule 17.1(a) or (b)

PCT/GB2005/000752

The Patent Office
Concept House
Cardiff Road
Newport
South Wales
NP10 8QQ

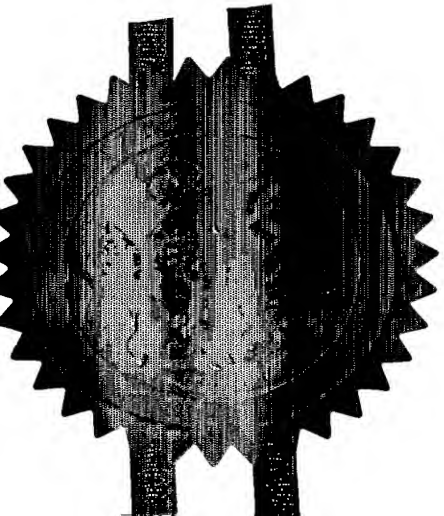
I, the undersigned, being an officer duly authorised in accordance with Section 74(1) and (4) of the Deregulation & Contracting Out Act 1994, to sign and issue certificates on behalf of the Comptroller-General, hereby certify that annexed hereto is a true copy of the documents as originally filed in connection with the patent application identified therein.

I also certify that the attached copy of the request for grant of a Patent (Form 1/77) bears an amendment, effected by this office, following a request by the applicant and agreed to by the Comptroller-General.

In accordance with the Patents (Companies Re-registration) Rules 1982, if a company named in this certificate and any accompanying documents has re-registered under the Companies Act 1980 with the same name as that with which it was registered immediately before re-registration save for the substitution as, or inclusion as, the last part of the name of the words "public limited company" or their equivalents in Welsh, references to the name of the company in this certificate and any accompanying documents shall be treated as references to the name with which it is so re-registered.

In accordance with the rules, the words "public limited company" may be replaced by p.l.c., plc, P.L.C. or PLC.

Re-registration under the Companies Act does not constitute a new legal entity but merely subjects the company to certain additional company law rules.



Signed A M Brewer

Dated 18 March 2005

Patents Form 1/77

Patents Act 1977
(Rule 16)

02MAR04 E877381-1 D00060
P01/7700 0.00-0404574.6 NONE

THE PATENT OFFICE
SN
- 1 MAR 2004
LONDON

Request for grant of a patent

(See the notes on the back of this form. You can also get an explanatory leaflet from the Patent Office to help you fill in this form)

The Patent Office

Cardiff Road
Newport
South Wales
NP10 8QQ

1. Your reference

RJW/LP6210215

2. Patent application number
(The Patent Office will fill this part in)

0404574.6

0 1 MAR 2004

3. Full name, address and postcode of the or of each applicant *(underline all surnames)*

SPIROGEN LIMITED
79 George Street
Ryde
ISLE OF WIGHT
PO33 2JF

Patents ADP number *(if you know it)*

8051872001

If the applicant is a corporate body, give the country/state of its incorporation

England

4. Title of the invention

AMINO ACIDS

5. Name of your agent *(if you have one)*

"Address for service" in the United Kingdom to which all correspondence should be sent *(including the postcode)*

MEWBURN ELLIS LLP
York House
23 Kingsway
London WC2B 6HP

Patents ADP number *(if you know it)*

109006

6. Priority: Complete this section if you are declaring priority from one or more earlier patent applications, filed in the last 12 months.

Country	Priority application number *(if you know it)*	Date of filing *(day / month / year)*

7. Divisionals, etc: Complete this section only if this application is a divisional application or resulted from an entitlement dispute (see note f)

Number of earlier UK application	Date of filing *(day / month / year)*

8. Is a Patents Form 7/77 (Statement of inventorship and of right to grant of a patent) required in support of this request?

YES

Answer YES if:

a) any applicant named in part 3 is not an inventor, or
b) there is an inventor who is not named as an applicant, or
c) any named applicant is a corporate body.

Otherwise answer NO (See note d)

Patents Form 1/77

9. Accompanying documents: A patent application must include a description of the invention. Not counting duplicates, please enter the number of pages of each item accompanying this form:

Continuation sheets of this form	
Description	51
Claim(s)	
Abstract	SN
Drawing(s)	1 X 1

10. If you are also filing any of the following, state how many against each item.

Priority documents	
Translations of priority documents	
Statement of inventorship and right to grant of a patent (Patents Form 7/77)	
Request for a preliminary examination and search (Patents Form 9/77)	
Request for a substantive examination (Patents Form 10/77)	
Any other documents (please specify)	

11. I/We request the grant of a patent on the basis of this application.

Signature(s) Mark Ellis Date 1 March 2004

12. Name, daytime telephone number and e-mail address, if any, of person to contact in the United Kingdom

ROBERT WATSON
020 7240 4405

Warning

After an application for a patent has been filed, the Comptroller of the Patent Office will consider whether publication or communication of the invention should be prohibited or restricted under Section 22 of the Patents Act 1977. You will be informed if it is necessary to prohibit or restrict your invention in this way. Furthermore, if you live in the United Kingdom, Section 23 of the Patents Act 1977 stops you from applying for a patent abroad without first getting written permission from the Patent Office unless an application has been filed at least 6 weeks beforehand in the United Kingdom for a patent for the same invention and either no direction prohibiting publication or communication has been given, or any such direction has been revoked.

Notes

a) If you need help to fill in this form or you have any questions, please contact the Patent Office on 08459 500505.

b) Write your answers in capital letters using black ink or you may type them.

c) If there is not enough space for all the relevant details on any part of this form, please continue on a separate sheet of paper and write "see continuation sheet" in the relevant part(s). Any continuation sheet should be attached to this form.

d) If you have answered YES in part 8, a Patents Form 7/77 will need to be filed.

e) Once you have filled in the form you must remember to sign and date it.

f) Part 7 should only be completed when a divisional application is being made under section 15(4), or when an application is being made under section 8(3), 12(6) or 37(4) following an entitlement dispute. By completing part 7 you are requesting that this application takes the same filing date as an earlier UK application. If you want the new application to have the same priority date(s) as the earlier UK application, you should also complete part 6 with the priority details.

AMINO ACIDS

The present invention relates to amino acids, i.e. compounds bearing amino and carboxy groups, their synthesis and use in synthesizing molecules designed to interact with DNA.

Background to the invention

The prototype minor groove binding agent distamycin A is a natural product with an amide linked tris(N-methylpyrrole) structure. The molecule binds non-covalently at A/T rich sequences and forms specific hydrogen bonds with the minor groove floor. The A/T recognising capacity of the molecule relates partly to favourable Van der Waals interactions with the groove walls in the relatively narrow A/T regions and also to specific steric clashes between the inner facing pyrrole H-3 and the larger G residues in the minor groove. The observation that distamycin and the related natural product netropsin may bind as highly cooperative 2:1 complexes in the minor groove was significant and prompted the development of a series of linked dimer molecules termed 'hairpin polyamides' (see for example, Woods, C.R., *et al.*, *J. Am. Chem. Soc.*, **124**, 10676-10682 (2002)) In such molecules replacement of the pyrrole with the sterically less demanding imidazole allows passive G recognition. A further development was the inclusion of a hydroxypyrrole unit which discriminated between T and A residues. Thus the full sequence recognising code is:

Heterocycle		Nucleotide	
Py	Py	A or T	A or T
Py	Hp	A	T
Hp	Py	T	A
Im	Py	G	C
Py	Im	C	G

For molecules which bind in a 1:1 motif with DNA the recognition properties are more degenerate, thus:

Heterocycle	Nucleotide
Py	A or T
Im	A or T or G or C
Hp	T?

More recently new heterocycles have been studied such as 2-(pyrrol-2-yl)benzimidazoles, 2-(pyrrol-2-yl)imidazopyridines and 5-hydroxy-(pyrrol-2-yl)benzimidazoles which have similar recognition properties to the established building blocks in the context of hairpin polyamides (Biehen, C.A., *et al.*, *Chem. Eur. J.*, **9**, 2110-2122 (2003)).

Disclosure of the invention

The present inventors have developed a series of compounds bearing amino and carboxy groups, which can be used in synthesising molecules designed to interact with DNA.

In a first aspect, the invention provides a compound of formula **I**:

Z′-CO-A-B-NH-Z (I)

wherein:

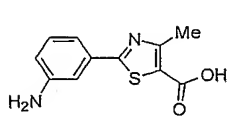
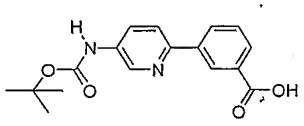
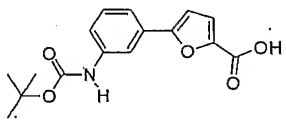
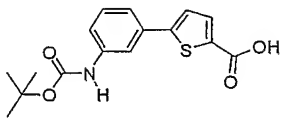
Z is H or an amino protecting group;

Z′ is OH, a protected or activated hydroxyl group or Cl;

A is an optionally substituted C_{5-6} arylene group;

B is an optionally substituted C_{5-6} arylene group;

except for the following compounds:

In a second aspect, the invention provides a method of synthesising a compound of formula **I**.

In a third aspect, the invention provides a polyamido moiety comprising at least one unit of formula II:

$$-CO-A-B-NH- \quad (II)$$

wherein:
A and B are as defined in the first aspect of the invention.

The unit of formula II may be bound to one or more other units selected from:
(i) units of formula II; and
(ii) amino-heteroarylene-carbonyl units of formula III:

$$-CO-E-NH- \quad (III)$$

wherein E is either optionally substituted C_{5-20} heteroarylene (G) or C_{8-10} heteroarylene-C_{5-20} arylene (K).

In a fourth aspect, the present invention provides the use of a compound of formula I as defined in the first aspect of the invention in the synthesis of a compound comprising a polyamido moiety as defined in the third aspect of the invention.

In a fifth aspect, the present invention provides a compound comprising a polyamido moiety as defined in the third aspect of the invention.

In some embodiments of the fifth aspect of the invention, the compounds are of formula IV:

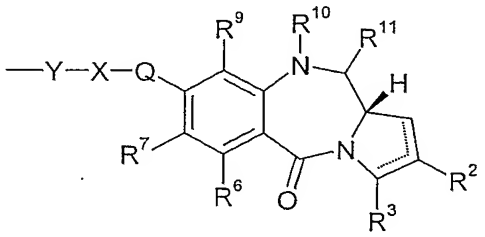
$$Z''-(T)_n-CO-(CH_2)_q-NR^1R^2 \quad (IV)$$

wherein:
Z" is OH or a protected hydroxy group;
each T is independently selected from units of formulae II, III or V:

$-CO-(CH_2)_{q'}-NH-$ (V)

wherein q' is from 1 to 3;
n is from 1 to 10;
q is from 1 to 3; and
R^1 and R^2 are independently selected from C_{1-4} alkyl.

In other embodiments of the fifth aspect of the invention, the compounds include a pyrrolobenzodiazpine moiety of formula **VI**:



(VI)

wherein:
the dotted lines indicate the optional presence of a double bond between C1 and C2 or C2 and C3;
R^2 and R^3 are independently selected from -H, -OH, =O, $=CH_2$, -CN, -R, OR, halo, =CH-R, $O-SO_2-R$, CO_2R and COR;
R^6, R^7 and R^9 are independently selected from H, R, OH, OR, SH, SR, NH_2, NHR, NHRR', nitro, Me_3Sn and halo;
where R and R' are independently selected from optionally substituted C_{1-7} alkyl, C_{3-20} heterocyclyl and C_{5-20} aryl groups; or
R^6 and R^7 together form a group $-O-(CH_2)_p-O-$, where p is 1 or 2;
R^{10} is a nitrogen protecting group and R^{11} is either $O-R^{15}$, wherein R^{15} is a hydroxyl protecting group, or OH, or R^{10} and R^{11} together form a double bond between N10 and C11;
Q is selected from O, S, NH or a single bond;
X is a divalent group such that HY = R, or a single bond;
Y is either NH or C(=O).

Further aspects of the present invention relate to compounds of the fifth aspect of the invention and pharmaceutical salts

thereof, their use in methods of therapy (particularly in treating gene-based diseases), pharmaceutical compositions comprising these, and their use in the manufacture of a medicament for the treatment of a gene-based disease.

Definitions

C_{5-6} arylene groups

The term C_{5-6} arylene, as used herein, pertains to a divalent moiety obtained by removing two hydrogen atoms from aromatic ring atoms of an aromatic compound and 5 or 6 ring atoms.

The ring atoms may be all carbon atoms, as in "carboaryl groups". Examples of carboaryl groups include, but are not limited to, those derived from benzene (i.e. phenylene) (C_6).

Alternatively, the ring atoms may include one or more heteroatoms, as in "C_{5-6} heteroarylene groups". Examples of C_{5-6} heteroarylene groups include, but are not limited to, those derived from:

N_1: pyrrole (azole) (C_5), pyridine (azine) (C_6);

O_1: furan (oxole) (C_5);

S_1: thiophene (thiole) (C_5);

N_1O_1: oxazole (C_5), isoxazole (C_5), isoxazine (C_6);

N_2O_1: oxadiazole (furazan) (C_5);

N_3O_1: oxatriazole (C_5);

N_1S_1: thiazole (C_5), isothiazole (C_5);

N_2: imidazole (1,3-diazole) (C_5), pyrazole (1,2-diazole) (C_5), pyridazine (1,2-diazine) (C_6), pyrimidine (1,3-diazine) (C_6) (e.g., cytosine, thymine, uracil), pyrazine (1,4-diazine) (C_6);

N_3: triazole (C_5), triazine (C_6); and,

N_4: tetrazole (C_5).

C_{5-20} heteroarylene groups (G)

G is an optionally substituted C_{5-20} heteroarylene group, preferably a C_{5-16} heteroarylene group, more preferably a C_{5-10} heteroarylene group and even more preferably a C_{5-6} heteroarylene group.

Furthermore in a preferred embodiment, the G group is a five membered heteroarylene group.

The heteroarylene group may contain one or more heteroatoms and preferably contains one heteroatom. The heteroatoms in the heteroarylene group are independently chosen from N, O and S and are preferably N.

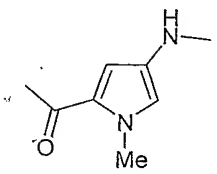
The heteroarylene G group is optionally substituted with one or more R groups. In a preferred embodiment the G group is substituted at one or more of the heteroatom positions with at least one R group, most preferably the R group is a methyl or ethyl group.

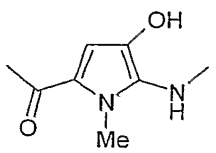
Where the G group is a six membered heteroarylene group, the -NH- and -CO- groups are preferably attached at the 2,6, 2,5, 3,6 or 3,5 positions.

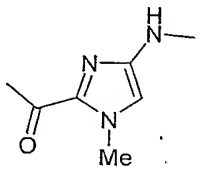
Where the G group is a five membered heteroarylene group, the -NH- and -CO- groups are preferably attached at the 2,5, 2,4 or 3,5 positions.

Where the G group comprises two fused rings, the -NH- and -CO- groups are preferably attached to different rings.

Known amino-C_5 heteroarylene-carbonyl units include the following:





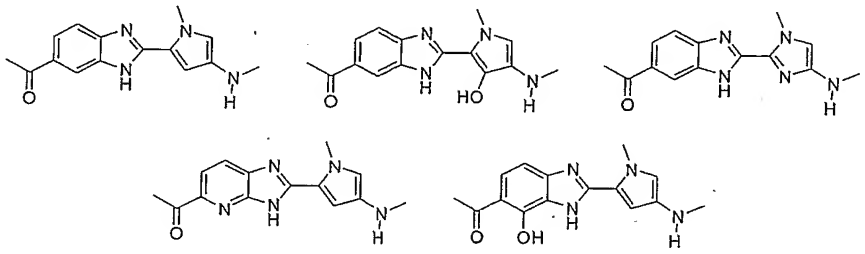


C_{8-10} heteroarylene-C_{5-20} arylene groups (K)

The C_{8-10} heteroarylene groups are a subset of the C_{5-20} heteroarylene groups defined above, and comprise two fused rings.

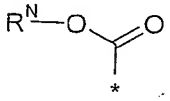
The term arylene, as used herein, pertains to a divalent moiety obtained by removing two hydrogen atoms from aromatic ring atoms of an aromatic compound having from 5 to 20 ring atoms. Arylene compounds as described herein correspond to aryl groups as defined below with one fewer hydrogen atoms on the ring atoms. Preferably, the C_{5-20} arylene group is a C_{5-7} arylene group and more preferably a C_{5-6} heteroarylene group.

Units of formula **III** which are a carbonyl-C_{8-10} heteroarylene-C_{5-6} heteroarylene-amino unit have been described in Briehen, C.A., *et al.*, *Chem. Eur. J.*, **9**, 2110-2122 (2003) and Renneberg, D., *et al.*, *J. Am. Chem. Soc.*, **125**, 5707-5716 (2003) and include:



Amino protecting groups (Z)

Amino protecting groups are well known in the art, and are listed on pages 494 to 572 of Greene, T.W. and Wuts, G.M., Protective Groups in Organic Synthesis, 3rd Edition, John Wiley & Sons, Inc., 1999, which is incorporated herein by reference. Preferred nitrogen protecting groups are carbamate protecting groups that have the general formula:



Particularly preferred protecting groups include Alloc, Troc, Teoc, Boc, and Fmoc, with Boc being particularly preferred.

Protected hydroxyl groups (Z')

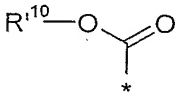
Protected hydroxyl groups are of the formula -O-Prot^{O}, where Prot^{O} is an oxygen protecting group as discussed below.

Activated hydroxyl groups (Z')

Activated hydroxyl groups are of the formula -O-Act, where Act is an activating moiety for peptide bond formation, introduced by a peptide coupling reagent. Such reagents include BOP, BOP-Cl, DCC, DIC, EDPP, HATU, HOBt, PyBroP and TBTU.

Nitrogen protecting groups (R^{10})

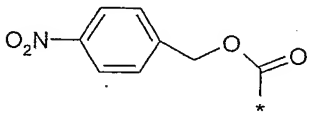
Nitrogen protecting groups are well known in the art. Preferred nitrogen protecting groups are carbamate protecting groups that have the general formula:



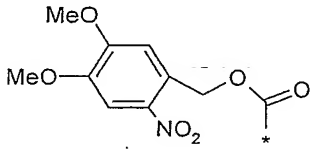
A large number of possible carbamate nitrogen protecting groups are listed on pages 503 to 549 of Greene, T.W. and Wuts, G.M., Protective Groups in Organic Synthesis, 3rd Edition, John Wiley & Sons, Inc., 1999, which is incorporated herein by reference.

Particularly preferred protecting groups include Alloc, Troc, Teoc, BOC, Doc, Hoc, TcBOC, Fmoc, 1-Adoc and 2-Adoc.

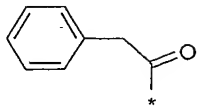
Also suitable for use in the present invention are nitrogen protecting groups which can be removed *in vivo* (e.g. enzymatically, using light) as described in WO 00/12507, which is incorporated herein by reference. Examples of these protecting groups include:



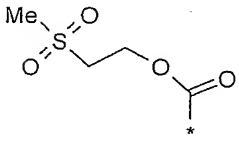
, which is nitroreductase labile (e.g. using ADEPT/GDEPT);



and



, which are photolabile; and



which is glutathione labile (e.g. using NPEPT).

Oxygen protecting groups

Oxygen protecting groups are well known in the art. A large number of suitable groups are described on pages 23 to 200 of Greene, T.W. and Wuts, G.M., Protective Groups in Organic Synthesis, 3rd Edition, John Wiley & Sons, Inc., 1999, which is incorporated herein by reference.

Classes of particular interest include silyl ethers, methyl ethers, alkyl ethers, benzyl ethers, esters, benzoates, carbonates, and sulfonates.

Substituents

The phrase "optionally substituted" as used herein, pertains to a parent group which may be unsubstituted or which may be substituted.

Unless otherwise specified, the term "substituted" as used herein, pertains to a parent group which bears one or more substitutents. The term "substituent" is used herein in the conventional sense and refers to a chemical moiety which is covalently attached to, or if appropriate, fused to, a parent group. A wide variety of substituents are well known, and methods for their formation and introduction into a variety of parent groups are also well known.

Examples of substituents are described in more detail below.

C_{1-7} alkyl: The term "C_{1-7} alkyl" as used herein, pertains to a monovalent moiety obtained by removing a hydrogen atom from a carbon atom of a hydrocarbon compound having from 1 to 7 carbon atoms, which may be aliphatic or alicyclic, and which may be saturated or unsaturated (e.g. partially unsaturated, fully unsaturated). Thus, the term "alkyl" includes the sub-classes alkenyl, alkynyl, cycloalkyl, etc., discussed below.

Examples of saturated alkyl groups include, but are not limited to, methyl (C_1), ethyl (C_2), propyl (C_3), butyl (C_4), pentyl (C_5), hexyl (C_6) and heptyl (C_7).

Examples of saturated linear alkyl groups include, but are not limited to, methyl (C_1), ethyl (C_2), n-propyl (C_3), n-butyl (C_4), n-pentyl (amyl) (C_5), n-hexyl (C_6) and n-heptyl (C_7).

Examples of saturated branched alkyl groups include iso-propyl (C_3), iso-butyl (C_4), sec-butyl (C_4), tert-butyl (C_4), iso-pentyl (C_5), and neo-pentyl (C_5).

C_{2-7} Alkenyl: The term "C_{2-7} alkenyl" as used herein, pertains to an alkyl group having one or more carbon-carbon double bonds.

Examples of unsaturated alkenyl groups include, but are not limited to, ethenyl (vinyl, $-CH=CH_2$), 1-propenyl ($-CH=CH-CH_3$), 2-propenyl (allyl, $-CH-CH=CH_2$), isopropenyl (1-methylvinyl, $-C(CH_3)=CH_2$), butenyl (C_4), pentenyl (C_5), and hexenyl (C_6).

C_{2-7} alkynyl: The term "C_{2-7} alkynyl" as used herein, pertains to an alkyl group having one or more carbon-carbon triple bonds.

Examples of unsaturated alkynyl groups include, but are not limited to, ethynyl (ethinyl, $-C\equiv CH$) and 2-propynyl (propargyl, $-CH_2-C\equiv CH$).

C_{3-7} cycloalkyl: The term "C_{3-7} cycloalkyl" as used herein, pertains to an alkyl group which is also a cyclyl group; that is, a monovalent moiety obtained by removing a hydrogen atom from an alicyclic ring atom of a cyclic hydrocarbon (carbocyclic) compound, which moiety has from 3 to 7 carbon atoms, including from 3 to 7 ring atoms.

Examples of cycloalkyl groups include, but are not limited to, those derived from:

saturated monocyclic hydrocarbon compounds: cyclopropane (C_3), cyclobutane (C_4), cyclopentane (C_5), cyclohexane (C_6), cycloheptane (C_7), methylcyclopropane (C_4), dimethylcyclopropane (C_5), methylcyclobutane (C_5), dimethylcyclobutane (C_6), methylcyclopentane (C_6), dimethylcyclopentane (C_7) and methylcyclohexane (C_7);

unsaturated monocyclic hydrocarbon compounds: cyclopropene (C_3), cyclobutene (C_4), cyclopentene (C_5), cyclohexene (C_6), methylcyclopropene (C_4), dimethylcyclopropene (C_5), methylcyclobutene (C_5), dimethylcyclobutene (C_6), methylcyclopentene (C_6), dimethylcyclopentene (C_7) and methylcyclohexene (C_7); and

saturated polycyclic hydrocarbon compounds: norcarane (C_7), norpinane (C_7), norbornane (C_7).

C_{3-20} heterocyclyl: The term "C_{3-20} heterocyclyl" as used herein, pertains to a monovalent moiety obtained by removing a hydrogen atom from a ring atom of a heterocyclic compound, which moiety has from 3 to 20 ring atoms, of which from 1 to 10 are ring heteroatoms. Preferably, each ring has from 3 to 7 ring atoms, of which from 1 to 4 are ring heteroatoms.

In this context, the prefixes (e.g. C_{3-20}, C_{3-7}, C_{5-6}, etc.) denote the number of ring atoms, or range of number of ring atoms, whether carbon atoms or heteroatoms. For example, the term "C_{5-6}heterocyclyl", as used herein, pertains to a heterocyclyl group having 5 or 6 ring atoms.

Examples of monocyclic heterocyclyl groups include, but are not limited to, those derived from:

N_1: aziridine (C_3), azetidine (C_4), pyrrolidine (tetrahydropyrrole) (C_5), pyrroline (e.g., 3-pyrroline, 2,5-dihydropyrrole) (C_5), 2H-pyrrole or 3H-pyrrole (isopyrrole, isoazole) (C_5), piperidine (C_6), dihydropyridine (C_6), tetrahydropyridine (C_6), azepine (C_7);

O_1: oxirane (C_3), oxetane (C_4), oxolane (tetrahydrofuran) (C_5), oxole (dihydrofuran) (C_5), oxane (tetrahydropyran) (C_6), dihydropyran (C_6), pyran (C_6), oxepin (C_7);

S_1: thiirane (C_3), thietane (C_4), thiolane (tetrahydrothiophene) (C_5), thiane (tetrahydrothiopyran) (C_6), thiepane (C_7);

O_2: dioxolane (C_5), dioxane (C_6), and dioxepane (C_7);

O_3: trioxane (C_6);

N_2: imidazolidine (C_5), pyrazolidine (diazolidine) (C_5), imidazoline (C_5), pyrazoline (dihydropyrazole) (C_5), piperazine (C_6);

N_1O_1: tetrahydrooxazole (C_5), dihydrooxazole (C_5), tetrahydroisoxazole (C_5), dihydroisoxazole (C_5), morpholine (C_6), tetrahydrooxazine (C_6), dihydrooxazine (C_6), oxazine (C_6);

N_1S_1: thiazoline (C_5), thiazolidine (C_5), thiomorpholine (C_6);

N_2O_1: oxadiazine (C_6);

O_1S_1: oxathiole (C_5) and oxathiane (thioxane) (C_6); and,

$N_1O_1S_1$: oxathiazine (C_6).

Examples of substituted monocyclic heterocyclyl groups include those derived from saccharides, in cyclic form, for example, furanoses (C_5), such as arabinofuranose, lyxofuranose, ribofuranose, and xylofuranse, and pyranoses (C_6), such as allopyranose, altropyranose, glucopyranose, mannopyranose, gulopyranose, idopyranose, galactopyranose, and talopyranose.

C_{5-20} aryl: The term "C_{5-20} aryl", as used herein, pertains to a monovalent moiety obtained by removing a hydrogen atom from an aromatic ring atom of an aromatic compound, which moiety has from

3 to 20 ring atoms. Preferably, each ring has from 5 to 7 ring atoms.

In this context, the prefixes (e.g. C_{3-20}, C_{5-7}, C_{5-6}, etc.) denote the number of ring atoms, or range of number of ring atoms, whether carbon atoms or heteroatoms. For example, the term "C_{5-6} aryl" as used herein, pertains to an aryl group having 5 or 6 ring atoms.

The ring atoms may be all carbon atoms, as in "carboaryl groups". Examples of carboaryl groups include, but are not limited to, those derived from benzene (i.e. phenyl) (C_6), naphthalene (C_{10}), azulene (C_{10}), anthracene (C_{14}), phenanthrene (C_{14}), naphthacene (C_{18}), and pyrene (C_{16}).

Examples of aryl groups which comprise fused rings, at least one of which is an aromatic ring, include, but are not limited to, groups derived from indane (e.g. 2,3-dihydro-1H-indene) (C_9), indene (C_9), isoindene (C_9), tetraline (1,2,3,4-tetrahydronaphthalene (C_{10}), acenaphthene (C_{12}), fluorene (C_{13}), phenalene (C_{13}), acephenanthrene (C_{15}), and aceanthrene (C_{16}).

Alternatively, the ring atoms may include one or more heteroatoms, as in "heteroaryl groups". Examples of monocyclic heteroaryl groups include, but are not limited to, those derived from:

N_1: pyrrole (azole) (C_5), pyridine (azine) (C_6);

O_1: furan (oxole) (C_5);

S_1: thiophene (thiole) (C_5);

N_1O_1: oxazole (C_5), isoxazole (C_5), isoxazine (C_6);

N_2O_1: oxadiazole (furazan) (C_5);

N_3O_1: oxatriazole (C_5);

N_1S_1: thiazole (C_5), isothiazole (C_5);

N_2: imidazole (1,3-diazole) (C_5), pyrazole (1,2-diazole) (C_5), pyridazine (1,2-diazine) (C_6), pyrimidine (1,3-diazine) (C_6) (e.g., cytosine, thymine, uracil), pyrazine (1,4-diazine) (C_6);

N_3: triazole (C_5), triazine (C_6); and,

N_4: tetrazole (C_5).

Examples of heteroaryl which comprise fused rings, include, but are not limited to:

C_9 (with 2 fused rings) derived from benzofuran (O_1), isobenzofuran (O_1), indole (N_1), isoindole (N_1), indolizine (N_1), indoline (N_1), isoindoline (N_1), purine (N_4) (e.g., adenine, guanine), benzimidazole (N_2), indazole (N_2), benzoxazole (N_1O_1), benzisoxazole (N_1O_1), benzodioxole (O_2), benzofurazan (N_2O_1), benzotriazole (N_3), benzothiofuran (S_1), benzothiazole (N_1S_1), benzothiadiazole (N_2S);

C_{10} (with 2 fused rings) derived from chromene (O_1), isochromene (O_1), chroman (O_1), isochroman (O_1), benzodioxan (O_2), quinoline (N_1), isoquinoline (N_1), quinolizine (N_1), benzoxazine (N_1O_1), benzodiazine (N_2), pyridopyridine (N_2), quinoxaline (N_2), quinazoline (N_2), cinnoline (N_2), phthalazine (N_2), naphthyridine (N_2), pteridine (N_4);

C_{11} (with 2 fused rings) derived from benzodiazepine (N_2);

C_{13} (with 3 fused rings) derived from carbazole (N_1), dibenzofuran (O_1), dibenzothiophene (S_1), carboline (N_2), perimidine (N_2), pyridoindole (N_2); and,

C_{14} (with 3 fused rings) derived from acridine (N_1), xanthene (O_1), thioxanthene (S_1), oxanthrene (O_2), phenoxathiin (O_1S_1), phenazine (N_2), phenoxazine (N_1O_1), phenothiazine (N_1S_1), thianthrene (S_2), phenanthridine (N_1), phenanthroline (N_2), phenazine (N_2).

The above groups, whether alone or part of another substituent, may themselves optionally be substituted with one or more groups selected from themselves and the additional substituents listed below.

Halo: -F, -Cl, -Br, and -I.

Hydroxy: -OH.

Ether: -OR, wherein R is an ether substituent, for example, a C_{1-7} alkyl group (also referred to as a C_{1-7} alkoxy group, discussed below), a C_{3-20} heterocyclyl group (also referred to as a C_{3-20} heterocyclyloxy group), or a C_{5-20} aryl group (also referred to as a C_{5-20} aryloxy group), preferably a C_{1-7}alkyl group.

Alkoxy: -OR, wherein R is an alkyl group, for example, a C_{1-7} alkyl group. Examples of C_{1-7} alkoxy groups include, but are not limited to, -OMe (methoxy), -OEt (ethoxy), -O(nPr) (n-propoxy), -O(iPr) (isopropoxy), -O(nBu) (n-butoxy), -O(sBu) (sec-butoxy), -O(iBu) (isobutoxy), and -O(tBu) (tert-butoxy).

Acetal: $-CH(OR^1)(OR^2)$, wherein R^1 and R^2 are independently acetal substituents, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group, or, in the case of a "cyclic" acetal group, R^1 and R^2, taken together with the two oxygen atoms to which they are attached, and the carbon atoms to which they are attached, form a heterocyclic ring having from 4 to 8 ring atoms. Examples of acetal groups include, but are not limited to, $-CH(OMe)_2$, $-CH(OEt)_2$, and -CH(OMe)(OEt).

Hemiacetal: $-CH(OH)(OR^1)$, wherein R^1 is a hemiacetal substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of hemiacetal groups include, but are not limited to, -CH(OH)(OMe) and -CH(OH)(OEt).

Ketal: $-CR(OR^1)(OR^2)$, where R^1 and R^2 are as defined for acetals, and R is a ketal substituent other than hydrogen, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples ketal groups include, but are not limited to, $-C(Me)(OMe)_2$, $-C(Me)(OEt)_2$, -C(Me)(OMe)(OEt), $-C(Et)(OMe)_2$, $-C(Et)(OEt)_2$, and -C(Et)(OMe)(OEt).

Hemiketal: $-CR(OH)(OR^1)$, where R^1 is as defined for hemiacetals, and R is a hemiketal substituent other than hydrogen, for example,

a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of hemiacetal groups include, but are not limited to, -C(Me)(OH)(OMe), -C(Et)(OH)(OMe), -C(Me)(OH)(OEt), and -C(Et)(OH)(OEt).

Oxo (keto, -one): =O.

Thione (thioketone): =S.

Imino (imine): =NR, wherein R is an imino substituent, for example, hydrogen, C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably hydrogen or a C_{1-7} alkyl group. Examples of ester groups include, but are not limited to, =NH, =NMe, =NEt, and =NPh.

Formyl (carbaldehyde, carboxaldehyde): -C(=O)H.

Acyl (keto): -C(=O)R, wherein R is an acyl substituent, for example, a C_{1-7} alkyl group (also referred to as C_{1-7} alkylacyl or C_{1-7} alkanoyl), a C_{3-20} heterocyclyl group (also referred to as C_{3-20} heterocyclylacyl), or a C_{5-20} aryl group (also referred to as C_{5-20} arylacyl), preferably a C_{1-7} alkyl group. Examples of acyl groups include, but are not limited to, $-C(=O)CH_3$ (acetyl), $-C(=O)CH_2CH_3$ (propionyl), $-C(=O)C(CH_3)_3$ (t-butyryl), and -C(=O)Ph (benzoyl, phenone).

Carboxy (carboxylic acid): -C(=O)OH.

Thiocarboxy (thiocarboxylic acid): -C(=S)SH.

Thiolocarboxy (thiolocarboxylic acid): -C(=O)SH.

Thionocarboxy (thionocarboxylic acid): -C(=S)OH.

Imidic acid: -C(=NH)OH.

Hydroxamic acid: -C(=NOH)OH.

Ester (carboxylate, carboxylic acid ester, oxycarbonyl): -C(=O)OR, wherein R is an ester substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of ester groups include, but are not limited to, $-C(=O)OCH_3$, $-C(=O)OCH_2CH_3$, $-C(=O)OC(CH_3)_3$, and -C(=O)OPh.

Acyloxy (reverse ester): -OC(=O)R, wherein R is an acyloxy substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of acyloxy groups include, but are not limited to, $-OC(=O)CH_3$ (acetoxy), $-OC(=O)CH_2CH_3$, $-OC(=O)C(CH_3)_3$, -OC(=O)Ph, and $-OC(=O)CH_2Ph$.

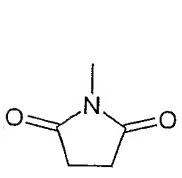
Oxycarboyloxy: -OC(=O)OR, wherein R is an ester substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of ester groups include, but are not limited to, $-OC(=O)OCH_3$, $-OC(=O)OCH_2CH_3$, $-OC(=O)OC(CH_3)_3$, and -OC(=O)OPh.

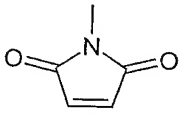
Amino: $-NR^1R^2$, wherein R^1 and R^2 are independently amino substituents, for example, hydrogen, a C_{1-7} alkyl group (also referred to as C_{1-7} alkylamino or di-C_{1-7} alkylamino), a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably H or a C_{1-7} alkyl group, or, in the case of a "cyclic" amino group, R^1 and R^2, taken together with the nitrogen atom to which they are attached, form a heterocyclic ring having from 4 to 8 ring atoms. Amino groups may be primary ($-NH_2$), secondary ($-NHR^1$), or tertiary ($-NHR^1R^2$), and in cationic form, may be quaternary ($-^+NR^1R^2R^3$). Examples of amino groups include, but are not limited to, $-NH_2$, $-NHCH_3$, $-NHC(CH_3)_2$, $-N(CH_3)_2$, $-N(CH_2CH_3)_2$, and -NHPh. Examples of cyclic amino groups include, but are not limited to, aziridino, azetidino, pyrrolidino, piperidino, piperazino, morpholino, and thiomorpholino.

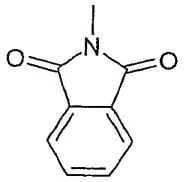
Amido (carbamoyl, carbamyl, aminocarbonyl, carboxamide): $-C(=O)NR^1R^2$, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of amido groups include, but are not limited to, $-C(=O)NH_2$, $-C(=O)NHCH_3$, $-C(=O)N(CH_3)_2$, $-C(=O)NHCH_2CH_3$, and $-C(=O)N(CH_2CH_3)_2$, as well as amido groups in which R^1 and R^2, together with the nitrogen atom to which they are attached, form a heterocyclic structure as in, for example, piperidinocarbonyl, morpholinocarbonyl, thiomorpholinocarbonyl, and piperazinocarbonyl.

Thioamido (thiocarbamyl): $-C(=S)NR^1R^2$, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of amido groups include, but are not limited to, $-C(=S)NH_2$, $-C(=S)NHCH_3$, $-C(=S)N(CH_3)_2$, and $-C(=S)NHCH_2CH_3$.

Acylamido (acylamino): $-NR^1C(=O)R^2$, wherein R^1 is an amide substituent, for example, hydrogen, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably hydrogen or a C_{1-7} alkyl group, and R^2 is an acyl substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20}aryl group, preferably hydrogen or a C_{1-7} alkyl group. Examples of acylamide groups include, but are not limited to, $-NHC(=O)CH_3$, $-NHC(=O)CH_2CH_3$, and $-NHC(=O)Ph$. R^1 and R^2 may together form a cyclic structure, as in, for example, succinimidyl, maleimidyl, and phthalimidyl:


succinimidyl


maleimidyl

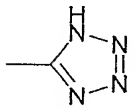

phthalimidyl

Aminocarbonyloxy: $-OC(=O)NR^1R^2$, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of aminocarbonyloxy groups include, but are not limited to, $-OC(=O)NH_2$, $-OC(=O)NHMe$, $-OC(=O)NMe_2$, and $-OC(=O)NEt_2$.

Ureido: $-N(R^1)CONR^2R^3$ wherein R^2 and R^3 are independently amino substituents, as defined for amino groups, and R^1 is a ureido substituent, for example, hydrogen, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably hydrogen or a C_{1-7} alkyl group. Examples of ureido groups include, but are not limited to, $-NHCONH_2$, -NHCONHMe, -NHCONHEt, $-NHCONMe_2$, $-NHCONEt_2$, $-NMeCONH_2$, -NMeCONHMe, -NMeCONHEt, $-NMeCONMe_2$, and $-NMeCONEt_2$.

Guanidino: $-NH-C(=NH)NH_2$.

Tetrazolyl: a five membered aromatic ring having four nitrogen atoms and one carbon atom,



Imino: =NR, wherein R is an imino substituent, for example, for example, hydrogen, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably H or a C_{1-7}alkyl group. Examples of imino groups include, but are not limited to, =NH, =NMe, and =NEt.

Amidine (amidino): $-C(=NR)NR_2$, wherein each R is an amidine substituent, for example, hydrogen, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably H or a C_{1-7} alkyl group. Examples of amidine groups include, but are not limited to, $-C(=NH)NH_2$, $-C(=NH)NMe_2$, and $-C(=NMe)NMe_2$.

Nitro: $-NO_2$.

Nitroso: -NO.

Azido: $-N_3$.

Cyano (nitrile, carbonitrile): -CN.

Isocyano: -NC.

Cyanato: -OCN.

Isocyanato: -NCO.

Thiocyano (thiocyanato): -SCN.

Isothiocyano (isothiocyanato): -NCS.

Sulfhydryl (thiol, mercapto): -SH.

Thioether (sulfide): -SR, wherein R is a thioether substituent, for example, a C_{1-7} alkyl group (also referred to as a C_{1-7}alkylthio group), a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of C_{1-7} alkylthio groups include, but are not limited to, $-SCH_3$ and $-SCH_2CH_3$.

Disulfide: -SS-R, wherein R is a disulfide substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group (also referred to herein as C_{1-7} alkyl disulfide). Examples of C_{1-7} alkyl disulfide groups include, but are not limited to, $-SSCH_3$ and $-SSCH_2CH_3$.

Sulfine (sulfinyl, sulfoxide): $-S(=O)R$, wherein R is a sulfine substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfine groups include, but are not limited to, $-S(=O)CH_3$ and $-S(=O)CH_2CH_3$.

Sulfone (sulfonyl): $-S(=O)_2R$, wherein R is a sulfone substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group, including, for example, a fluorinated or perfluorinated C_{1-7} alkyl group. Examples of sulfone groups include, but are not limited to, $-S(=O)_2CH_3$ (methanesulfonyl, mesyl), $-S(=O)_2CF_3$ (triflyl), $-S(=O)_2CH_2CH_3$

(esyl), $-S(=O)_2C_4F_9$ (nonaflyl), $-S(=O)_2CH_2CF_3$ (tresyl), $-S(=O)_2CH_2CH_2NH_2$ (tauryl), $-S(=O)_2Ph$ (phenylsulfonyl, besyl), 4-methylphenylsulfonyl (tosyl), 4-chlorophenylsulfonyl (closyl), 4-bromophenylsulfonyl (brosyl), 4-nitrophenyl (nosyl), 2-naphthalenesulfonate (napsyl), and 5-dimethylamino-naphthalen-1-ylsulfonate (dansyl).

Sulfinic acid (sulfino): $-S(=O)OH$, $-SO_2H$.

Sulfonic acid (sulfo): $-S(=O)_2OH$, $-SO_3H$.

Sulfinate (sulfinic acid ester): $-S(=O)OR$; wherein R is a sulfinate substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfinate groups include, but are not limited to, $-S(=O)OCH_3$ (methoxysulfinyl; methyl sulfinate) and $-S(=O)OCH_2CH_3$ (ethoxysulfinyl; ethyl sulfinate).

Sulfonate (sulfonic acid ester): $-S(=O)_2OR$, wherein R is a sulfonate substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfonate groups include, but are not limited to, $-S(=O)_2OCH_3$ (methoxysulfonyl; methyl sulfonate) and $-S(=O)_2OCH_2CH_3$ (ethoxysulfonyl; ethyl sulfonate).

Sulfinyloxy: $-OS(=O)R$, wherein R is a sulfinyloxy substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfinyloxy groups include, but are not limited to, $-OS(=O)CH_3$ and $-OS(=O)CH_2CH_3$.

Sulfonyloxy: $-OS(=O)_2R$, wherein R is a sulfonyloxy substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfonyloxy groups include, but are not limited to, $-OS(=O)_2CH_3$ (mesylate) and $-OS(=O)_2CH_2CH_3$ (esylate).

Sulfate: $-OS(=O)_2OR$; wherein R is a sulfate substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfate groups include, but are not limited to, $-OS(=O)_2OCH_3$ and $-SO(=O)_2OCH_2CH_3$.

Sulfamyl (sulfamoyl; sulfinic acid amide; sulfinamide): $-S(=O)NR^1R^2$, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of sulfamyl groups include, but are not limited to, $-S(=O)NH_2$, $-S(=O)NH(CH_3)$, $-S(=O)N(CH_3)_2$, $-S(=O)NH(CH_2CH_3)$, $-S(=O)N(CH_2CH_3)_2$, and $-S(=O)NHPh$.

Sulfonamido (sulfinamoyl; sulfonic acid amide; sulfonamide): $-S(=O)_2NR^1R^2$, wherein R^1 and R^2 are independently amino substituents, as defined for amino groups. Examples of sulfonamido groups include, but are not limited to, $-S(=O)_2NH_2$, $-S(=O)_2NH(CH_3)$, $-S(=O)_2N(CH_3)_2$, $-S(=O)_2NH(CH_2CH_3)$, $-S(=O)_2N(CH_2CH_3)_2$, and $-S(=O)_2NHPh$.

Sulfamino: $-NR^1S(=O)_2OH$, wherein R^1 is an amino substituent, as defined for amino groups. Examples of sulfamino groups include, but are not limited to, $-NHS(=O)_2OH$ and $-N(CH_3)S(=O)_2OH$.

Sulfonamino: $-NR^1S(=O)_2R$, wherein R^1 is an amino substituent, as defined for amino groups, and R is a sulfonamino substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfonamino groups include, but are not limited to, $-NHS(=O)_2CH_3$ and $-N(CH_3)S(=O)_2C_6H_5$.

Sulfinamino: $-NR^1S(=O)R$, wherein R^1 is an amino substituent, as defined for amino groups, and R is a sulfinamino substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group. Examples of sulfinamino

groups include, but are not limited to, $-NHS(=O)CH_3$ and $-N(CH_3)S(=O)C_6H_5$.

Phosphino (phosphine): $-PR_2$, wherein R is a phosphino substituent, for example, -H, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphino groups include, but are not limited to, $-PH_2$, $-P(CH_3)_2$, $-P(CH_2CH_3)_2$, $-P(t-Bu)_2$, and $-P(Ph)_2$.

Phospho: $-P(=O)_2$.

Phosphinyl (phosphine oxide): $-P(=O)R_2$, wherein R is a phosphinyl substituent, for example, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably a C_{1-7} alkyl group or a C_{5-20} aryl group. Examples of phosphinyl groups include, but are not limited to, $-P(=O)(CH_3)_2$, $-P(=O)(CH_2CH_3)_2$, $-P(=O)(t-Bu)_2$, and $-P(=O)(Ph)_2$.

Phosphonic acid (phosphono): $-P(=O)(OH)_2$.

Phosphonate (phosphono ester): $-P(=O)(OR)_2$, where R is a phosphonate substituent, for example, -H, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphonate groups include, but are not limited to, $-P(=O)(OCH_3)_2$, $-P(=O)(OCH_2CH_3)_2$, $-P(=O)(O-t-Bu)_2$, and $-P(=O)(OPh)_2$.

Phosphoric acid (phosphonooxy): $-OP(=O)(OH)_2$.

Phosphate (phosphonooxy ester): $-OP(=O)(OR)_2$, where R is a phosphate substituent, for example, -H, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphate groups include, but are not limited to, $-OP(=O)(OCH_3)_2$, $-OP(=O)(OCH_2CH_3)_2$, $-OP(=O)(O-t-Bu)_2$, and $-OP(=O)(OPh)_2$.

Phosphorous acid: $-OP(OH)_2$.

Phosphite: $-OP(OR)_2$, where R is a phosphite substituent, for example, -H, a C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphite groups include, but are not limited to, $-OP(OCH_3)_2$, $-OP(OCH_2CH_3)_2$, $-OP(O\text{-}t\text{-}Bu)_2$, and $-OP(OPh)_2$.

Phosphoramidite: $-OP(OR^1)-NR^2_2$, where R^1 and R^2 are phosphoramidite substituents, for example, -H, a (optionally substituted) C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphoramidite groups include, but are not limited to, $-OP(OCH_2CH_3)-N(CH_3)_2$, $-OP(OCH_2CH_3)-N(i\text{-}Pr)_2$, and $-OP(OCH_2CH_2CN)-N(i\text{-}Pr)_2$.

Phosphoramidate: $-OP(=O)(OR^1)-NR^2_2$, where R^1 and R^2 are phosphoramidate substituents, for example, -H, a (optionally substituted) C_{1-7} alkyl group, a C_{3-20} heterocyclyl group, or a C_{5-20} aryl group, preferably -H, a C_{1-7} alkyl group, or a C_{5-20} aryl group. Examples of phosphoramidate groups include, but are not limited to, $-OP(=O)(OCH_2CH_3)-N(CH_3)_2$, $-OP(=O)(OCH_2CH_3)-N(i\text{-}Pr)_2$, and $-OP(=O)(OCH_2CH_2CN)-N(i\text{-}Pr)_2$.

Gene-based diseases

Gene-based diseases include, and are preferably, proliferative diseases, and also include Alzheimer's disease and bacterial, parasitic and viral infections. Any condition which may be treated by the regulation of gene expression may be treated using compounds of the fifth aspect of the invention.

Proliferative Diseases

One of ordinary skill in the art is readily able to determine whether or not a candidate compound treats a proliferative condition for any particular cell type. For example, assays which

may conveniently be used to assess the activity offered by a particular compound are described in the examples below.

The term "proliferative disease" pertains to an unwanted or uncontrolled cellular proliferation of excessive or abnormal cells which is undesired, such as, neoplastic or hyperplastic growth, whether *in vitro* or *in vivo*.

Examples of proliferative conditions include, but are not limited to, benign, pre-malignant, and malignant cellular proliferation, including but not limited to, neoplasms and tumours (e.g. histocytoma, glioma, astrocyoma, osteoma), cancers (e.g. lung cancer, small cell lung cancer, gastrointestinal cancer, bowel cancer, colon cancer, breast carinoma, ovarian carcinoma, prostate cancer, testicular cancer, liver cancer, kidney cancer, bladder cancer, pancreas cancer, brain cancer, sarcoma, osteosarcoma, Kaposi's sarcoma, melanoma), leukemias, psoriasis, bone diseases, fibroproliferative disorders (e.g. of connective tissues), and atherosclerosis.

Any type of cell may be treated, including but not limited to, lung, gastrointestinal (including, e.g. bowel, colon), breast (mammary), ovarian, prostate, liver (hepatic), kidney (renal), bladder, pancreas, brain, and skin.

Methods of Treatment

As described above, the present invention provide the use of a compound of the fifth aspect in a method of therapy. If the compounds of the fifth aspect include a PBD moiety, then this preferably comprises a N10-C11 imine bond, or has a N10 which is protected by a nitrogen protecting group (R^{10}) which can be removed in vivo and the C11 substituent (R^{11}) as OH. Also provided is a method of treatment, comprising administering to a subject in need of treatment a therapeutically-effective amount of a compound of of the fifth aspect, preferably in the form of a pharmaceutical composition, which is the third aspect of the present invention.

The term "therapeutically effective amount" is an amount sufficient to show benefit to a patient. Such benefit may be at least amelioration of at least one symptom. The actual amount administered, and rate and time-course of administration, will depend on the nature and severity of what is being treated. Prescription of treatment, e.g. decisions on dosage, is within the responsibility of general practitioners and other medical doctors.

A compound may be administered alone or in combination with other treatments, either simultaneously or sequentially dependent upon the condition to be treated. Examples of treatments and therapies include, but are not limited to, chemotherapy (the administration of active agents, including, e.g. drugs; surgery; and radiation therapy. If the compound of formula of the fifth aspect comprises a PBD moiety which bears a carbamate-based nitrogen protecting group which may be removed *in vivo*, then the methods of treatment described in WO 00/12507 (ADEPT, GDEPT and PDT) may be used.

Pharmaceutical compositions according to the present invention, and for use in accordance with the present invention, may comprise, in addition to the active ingredient, i.e. a compound of formula of the fifth aspect, a pharmaceutically acceptable excipient, carrier, buffer, stabiliser or other materials well known to those skilled in the art. Such materials should be non-toxic and should not interfere with the efficacy of the active ingredient. The precise nature of the carrier or other material will depend on the route of administration, which may be oral, or by injection, e.g. cutaneous, subcutaneous, or intravenous.

Pharmaceutical compositions for oral administration may be in tablet, capsule, powder or liquid form. A tablet may comprise a solid carrier or an adjuvant. Liquid pharmaceutical compositions generally comprise a liquid carrier such as water, petroleum, animal or vegetable oils, mineral oil or synthetic oil. Physiological saline solution, dextrose or other saccharide solution or glycols such as ethylene glycol, propylene glycol or

polyethylene glycol may be included. A capsule may comprise a solid carrier such a gelatin.

For intravenous, cutaneous or subcutaneous injection, or injection at the site of affliction, the active ingredient will be in the form of a parenterally acceptable aqueous solution which is pyrogen-free and has suitable pH, isotonicity and stability. Those of relevant skill in the art are well able to prepare suitable solutions using, for example, isotonic vehicles such as Sodium Chloride Injection, Ringer's Injection, Lactated Ringer's Injection. Preservatives, stabilisers, buffers, antioxidants and/or other additives may be included, as required.

Includes Other Forms

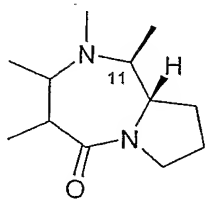
Unless otherwise specified, included in the above are the well known ionic, salt, solvate, and protected forms of these substituents. For example, a reference to carboxylic acid (-COOH) also includes the anionic (carboxylate) form ($-COO^-$), a salt or solvate thereof, as well as conventional protected forms. Similarly, a reference to an amino group includes the protonated form ($-N^+HR^1R^2$), a salt or solvate of the amino group, for example, a hydrochloride salt, as well as conventional protected forms of an amino group. Similarly, a reference to a hydroxyl group also includes the anionic form ($-O^-$), a salt or solvate thereof, as well as conventional protected forms.

Isomers, Salts and Solvates

Certain compounds may exist in one or more particular geometric, optical, enantiomeric, diasteriomeric, epimeric, atropic, stereoisomeric, tautomeric, conformational, or anomeric forms, including but not limited to, cis- and trans-forms; E- and Z-forms; c-, t-, and r- forms; endo- and exo-forms; R-, S-, and meso-forms; D- and L-forms; d- and l-forms; (+) and (-) forms; keto-, enol-, and enolate-forms; syn- and anti-forms; synclinal- and anticlinal-forms; α- and β-forms; axial and equatorial forms; boat-, chair-, twist-, envelope-, and halfchair-forms; and

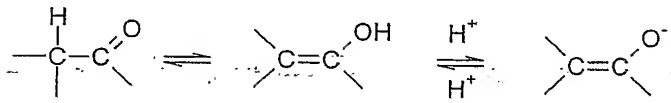
combinations thereof, hereinafter collectively referred to as "isomers" (or "isomeric forms").

Preferably if the compound of the fifth aspect comprise a PBD moiety then this moiety has the following stereochemistry at the C11 position:



Note that, except as discussed below for tautomeric forms, specifically excluded from the term "isomers", as used herein, are structural (or constitutional) isomers (i.e. isomers which differ in the connections between atoms rather than merely by the position of atoms in space). For example, a reference to a methoxy group, $-OCH_3$, is not to be construed as a reference to its structural isomer, a hydroxymethyl group, $-CH_2OH$. Similarly, a reference to ortho-chlorophenyl is not to be construed as a reference to its structural isomer, meta-chlorophenyl. However, a reference to a class of structures may well include structurally isomeric forms falling within that class (e.g. C_{1-7} alkyl includes n-propyl and iso-propyl; butyl includes n-, iso-, sec-, and tert-butyl; methoxyphenyl includes ortho-, meta-, and para-methoxyphenyl).

The above exclusion does not pertain to tautomeric forms, for example, keto-, enol-, and enolate-forms, as in, for example, the following tautomeric pairs: keto/enol (illustrated below), imine/enamine, amide/imino alcohol, amidine/amidine, nitroso/oxime, thioketone/enethiol, N-nitroso/hyroxyazo, and nitro/aci-nitro.



Note that specifically included in the term "isomer" are compounds with one or more isotopic substitutions. For example, H may be in any isotopic form, including ^{1}H, ^{2}H (D), and ^{3}H (T); C may be in any isotopic form, including ^{12}C, ^{13}C, and ^{14}C; O may be in any isotopic form, including ^{16}O and ^{18}O; and the like.

Unless otherwise specified, a reference to a particular compound includes all such isomeric forms, including (wholly or partially) racemic and other mixtures thereof. Methods for the preparation (e.g. asymmetric synthesis) and separation (e.g. fractional crystallisation and chromatographic means) of such isomeric forms are either known in the art or are readily obtained by adapting the methods taught herein, or known methods, in a known manner.

Unless otherwise specified, a reference to a particular compound also includes ionic, salt, solvate, and protected forms of thereof, for example, as discussed below.

It may be convenient or desirable to prepare, purify, and/or handle a corresponding salt of the active compound, for example, a pharmaceutically-acceptable salt. Examples of pharmaceutically acceptable salts are discussed in Berge, *et al.*, *J. Pharm. Sci.*, **66**, 1-19 (1977).

For example, if the compound is anionic, or has a functional group which may be anionic (e.g. -COOH may be $-COO^{-}$), then a salt may be formed with a suitable cation. Examples of suitable inorganic cations include, but are not limited to, alkali metal ions such as Na^{+} and K^{+}, alkaline earth cations such as Ca^{2+} and Mg^{2+}, and other cations such as Al^{+3}. Examples of suitable organic cations include, but are not limited to, ammonium ion (i.e. NH_4^{+}) and substituted ammonium ions (e.g. NH_3R^{+}, $NH_2R_2^{+}$, NHR_3^{+}, NR_4^{+}). Examples of some suitable substituted ammonium ions are those derived from: ethylamine, diethylamine, dicyclohexylamine, triethylamine, butylamine, ethylenediamine, ethanolamine,

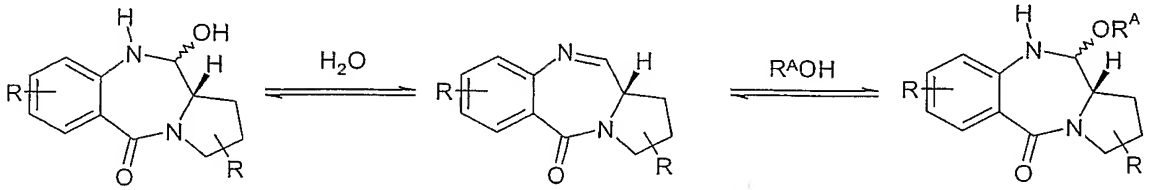
diethanolamine, piperazine, benzylamine, phenylbenzylamine, choline, meglumine, and tromethamine, as well as amino acids, such as lysine and arginine. An example of a common quaternary ammonium ion is $N(CH_3)_4^+$.

If the compound is cationic, or has a functional group which may be cationic (e.g. $-NH_2$ may be $-NH_3^+$), then a salt may be formed with a suitable anion. Examples of suitable inorganic anions include, but are not limited to, those derived from the following inorganic acids: hydrochloric, hydrobromic, hydroiodic, sulfuric, sulfurous, nitric, nitrous, phosphoric, and phosphorous.

Examples of suitable organic anions include, but are not limited to, those derived from the following organic acids: 2-acetyoxybenzoic, acetic, ascorbic, aspartic, benzoic, camphorsulfonic, cinnamic, citric, edetic, ethanedisulfonic, ethanesulfonic, fumaric, glucheptonic, gluconic, glutamic, glycolic, hydroxymaleic, hydroxynaphthalene carboxylic, isethionic, lactic, lactobionic, lauric, maleic, malic, methanesulfonic, mucic, oleic, oxalic, palmitic, pamoic, pantothenic, phenylacetic, phenylsulfonic, propionic, pyruvic, salicylic, stearic, succinic, sulfanilic, tartaric, toluenesulfonic, and valeric. Examples of suitable polymeric organic anions include, but are not limited to, those derived from the following polymeric acids: tannic acid, carboxymethyl cellulose.

It may be convenient or desirable to prepare, purify, and/or handle a corresponding solvate of the active compound. The term "solvate" is used herein in the conventional sense to refer to a complex of solute (e.g. active compound, salt of active compound) and solvent. If the solvent is water, the solvate may be conveniently referred to as a hydrate, for example, a mono-hydrate, a di-hydrate, a tri-hydrate, etc.

If the compounds of the fifth aspect comprise a PBD moiety then solvates of particular relevance are those where the solvent adds across the imine bond of the PBD moiety, which is illustrated below where the solvent is water or an alcohol (R^AOH, where R^A is an ether substituent as described above):



These forms can be called the carbinolamine and carbinolamine ether forms of the PBD. The balance of these equilibria depend on the conditions in which the compounds are found, as well as the nature of the moiety itself.

In general any nuceleophilic solvent is capable of forming such solvates as illustrated above for hydoxylic solvents. Other nuceleophilic solvents include thiols and amines.

These solvates may be isolated in solid form, for example, by lyophilisation.

General synthetic routes

Compounds of formula **I** may be made by a variety of routes, some of which are discussed below.

Modification of commercially available materials

Some compounds of formula 1 are commercially available:

HO_2C-A-B-NO_2 **Formula 1**

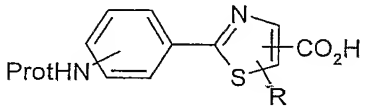
and can be readily modified to give compounds of formula **I**. The first step is protection of the carboxy group, for example as a methyl ester, using, for example EDCL, DMAP and MeOH in DMF. The nitro group can then be reduced to an amino group, for example using H_2 with a Pd/C catalyst in ethanol. The amino group can be protected, if necessary, for example by the use of $BocO_2$ in THF;

and the carboxy group may be deprotected by hydrolysis, for example using NaOH.

Heterocylic ring closures

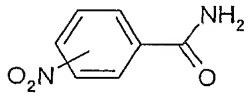
If one of A and B is a heteroarylene group, then the compound of formula I can be synthesised from an appropriate precursor by means of a ring closure reaction.

For example, compounds of formula 2:



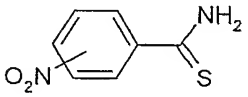
Formula 2

may be synthesised from a compound of formula 3:



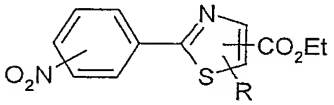
Formula 3

The compound of formula 3 can be converted into a compound of formula 4:



Formula 4

by using Lawessons regent, and this can then be ring closed to form a compounds of formula 5:



Formula 5

The ring closure may be accomplished by known methods, for example reaction with ethyl bromopyruvate or ethyl 2-chloroacetoaceatate. The nitro group may be converted to an amine group and protected in a similar manner to that discussed above, and the carboxy group may be deprotected also in a similar way to above.

Suzuki coupling

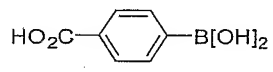
Compounds of formula I may be synthesised by the coupling of the two C_{5-6} arylene groups using Suzuki methodology. The groups -NHZ and -CO-Z' may be present on the C_{5-6} arylene groups in their final

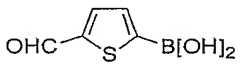
form prior to coupling, or may be present as precursors (for example, a precursor to -NHZ is NO_2 - see above for conversion; a precursor of CO-Z' is -CHO, which can be converted by oxidation and optionally protected).

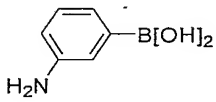
The coupling groups (e.g. Br and boronic acid/ester) may be either way round on A and B.

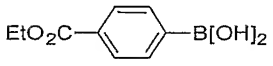
Suitable commercially available arylboronic acids/esters include:

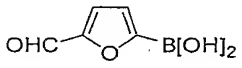
Boronic acids and Esters

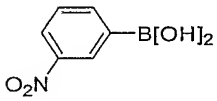


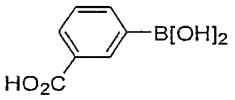


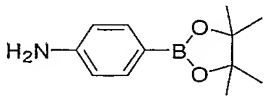


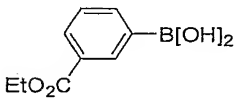






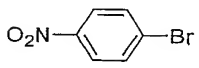
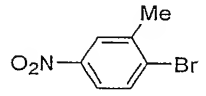
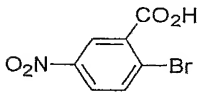
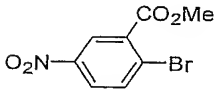
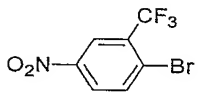
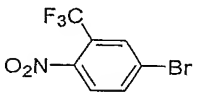
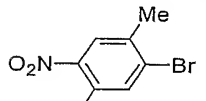
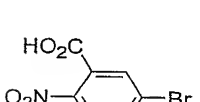


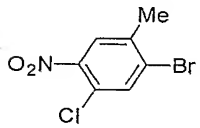




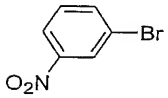
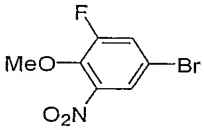
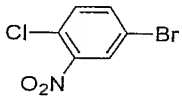
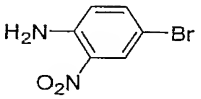
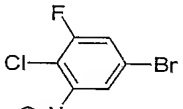
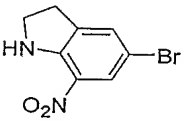
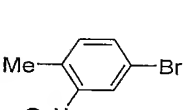
Suitable commercially available bromo compounds include:

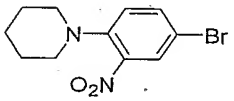
1-Bromo-4-nitrobenzenes

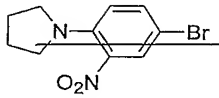
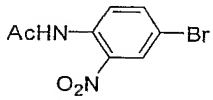
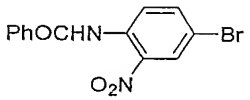
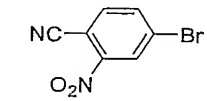
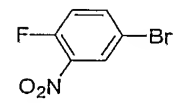
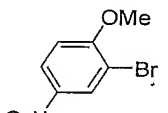











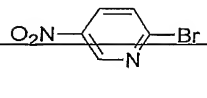
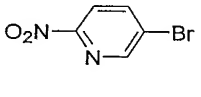
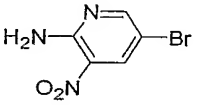
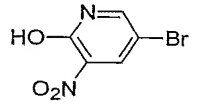
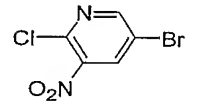
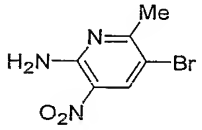
1-Bromo-3-nitrobenzenes

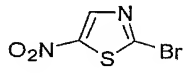
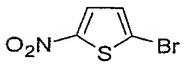


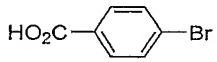
Pyridine Systems

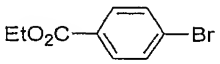







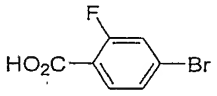
5-Membered Rings

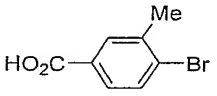



4-Bromobenzoic acids

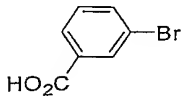


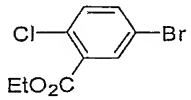


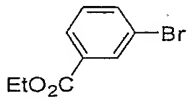


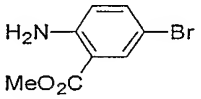


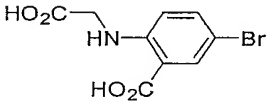
3-Bromobenzoic acids

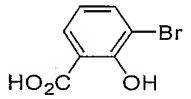


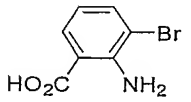


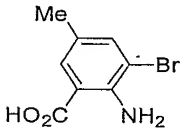


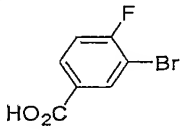


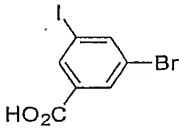


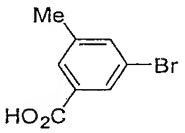


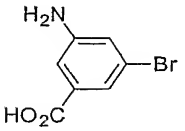


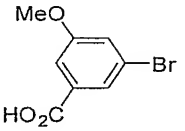




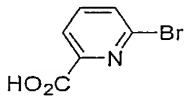


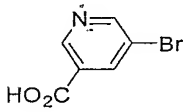


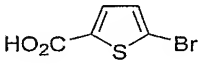


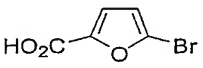


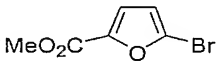
Heterocyclic bromocarboxylic acids

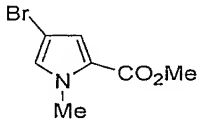


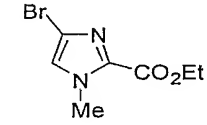












Polyamido moieties

Polyamido moieties comprising a unit of formula II may be synthesised by reacting a compound of formula I with a compound having an amino or carboxy (or equivalent) terminating group. Typically, one end of compound I will be protected to prevent self-condensation. The other units described in the second aspect

are well known, as are methods of amide bond formation. Typically, the carboxy group may be activated as an acid chloride group, or coupling initiators used, e.g. HOBt and EDCI.

Pyrrolobenzodiazepine moieties

The synthesis of pyrrolobenzodiazepine moieties as described in the fifth aspect of the invention are described in WO 00/12506. Protection at the C11 position can be readily introduced.

Compounds of formula IV

Compounds of formula **IV** may be synthesised by reacting a polyamido chain of formula 6:

$$Z''-(T)_n-Z' \quad \textbf{Formula 6}$$

with a compound of formula 7:

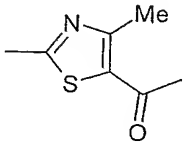
$$Z'''-CO-(CH_2)_q-NR^1R^2 \quad \textbf{Formula 7}$$

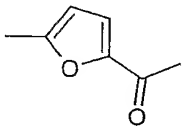
where Z''' is either OH or Cl, under amide bond forming conditions as described above. The compound of formula 6 may be formed as discussed above for polyamido moieties.

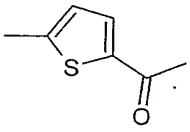
Further preferences

The following preferences may apply to all aspects of the invention as described above, or may relate to a single aspect. The preferences may be combined together in any combination.

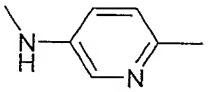
In the first aspect of the invention, it is preferred that if B is phenylene with -NH- β to the bond between A and B, then -A-CO- is not:



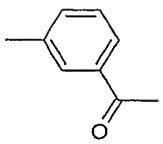




In the first aspect, it is also preferred that if -NH-B- is :



then -A-CO- is not:



In the first aspect, it is further preferred that if B is phenylene, then A is not thiazolylene, furanylene or thiphenylene and if B is pyridylene, then A is not phenylene.

If Z' is a protected hydroxyl group, it is preferably and alkoxy group, and more preferably methoxy or ethoxy.

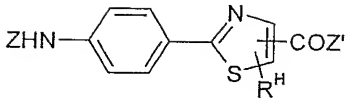
It is preferred that A and B are independently selected from phenylene, and arylene groups derived from C_5 heteroaryl groups having one or two heteroatoms, preferably at least one of which is nitrogen (i.e. pyrrole, oxazole, isoxzaole, thiazole, isothiazole, imidazole and pyrazole). Of these, pyrrole, oxazole, thiazole and imidazole are preferred. The nitrogen atom of these groups may be substituted with a C_{1-4} alkyl group, which is more preferably methyl.

The amino and carbonyl groups are preferably bound to A and B respectively at a position β- or γ- to the bond between A and B (i.e. not adjacent to the bond between A and B).

In some embodiments, one of A and B is phenylene and the other of A and B is a C_5-heteroarylene group, preferably with one or two hetero ring atoms, one of which ring atoms is nitrogen.

Preferable substituents on A and B include, but are not limited to: C_{1-4} alkyl (e.g. Me, CF_3), C_{1-4} alkoxy (e.g. MeO, EtO), halo (e.g. Cl, F) and amino, preferably susbtituted by one or two C_{1-4} alkyl groups.

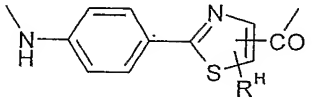
Particularly preferred compounds of formula I are of formula Ia:



Ia

where R^H is selected from H and C_{1-4} alkyl, and is preferably H or Me.

Particularly preferred units of formula **II** are of formula **IIa**:



IIa

where R^H is selected from H and C_{1-4} alkyl, and is preferably H or Me.

If compounds of the fifth aspect comprise a PBD moiety, then the following preferences are relevant:

R^9 is preferably H.

R^6 is preferably selected from H, OH, OR, SH, NH_2, nitro and halo, and is more preferably H or halo, and most preferably is H.

R^7 is preferably independently selected from H, OR, SH, SR, NH_2, NHR, NHRR', and halo, and more preferably independently selected from H and OR, where R is preferably selected from optionally substituted C_{1-7} alkyl, C_{3-10} heterocyclyl and C_{5-10} aryl groups. Preferably R^7 is OMe or H and most preferably OMe.

R^{10} is preferably BOC, Troc or alloc and is most preferably alloc.

R^{15} is preferably THP or a silyl oxygen protecting group (for example TBS) and is most preferably THP.

In other embodiments of the invention, R^{10} and R^{11} together form a double bond between N10 and C11.

Q is preferably NH, O or a single bond and most preferably NH or O.

X is preferably a single bond or C_{1-7} alkylene, more preferably a single bond or C_3 alkylene.

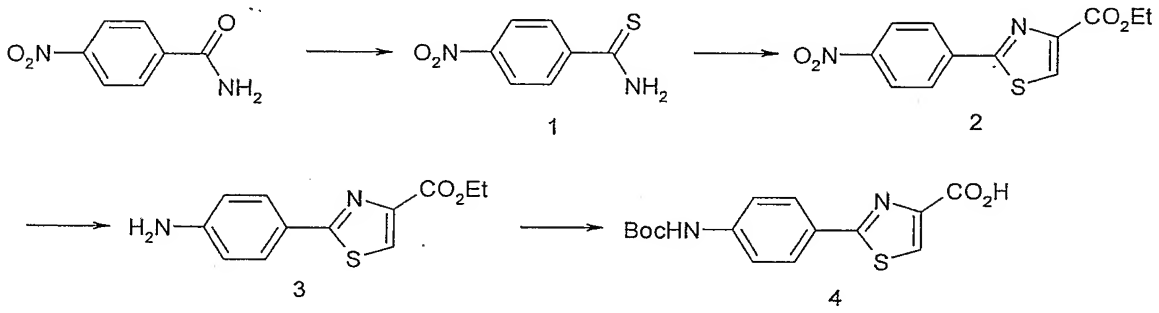
R^3 is preferably H.

R^2 is preferably R, and is more preferably an optionally substituted C_{5-20} aryl group. Most preferred is an optionally substituted phenyl group.

Brief Description of Figure

Figure 1 shows an illustrative example of an elctrophoresis gel in a DNA footprinting experiment (see Example 7).

Example 1 - 2-(4-tert-Butyloxycarbonylaminophenyl)thiazole-4-carboxylic acid (4)



(a) 4-Nitrothiobenzamide (1)

A suspension of 4-nitrobenzamide (5g, 30.1mmol) in chlorobenzene (150mL) was stirred at 80°C and Lawessons reagent (7.3g, 18.1mmol, 0.6equiv.) was added. The reaction mixture became orange/red in colour and all of the starting material dissolved. The solution was allowed to cool to room temperature and was stirred overnight. The precipitate formed was collected on a filter, washed with hexane then dried under vacuum. The crude product (6.0g) was recrystallised from ethanol/water to give the product 1 as gold coloured needles (3.81g, 70%).

^{1}H NMR (d_6-DMSO) δ 10.20 (1H, bs, N-H), 9.80 (1H, bs, N-H), 8.25 (2H, d, J = 8.9Hz, H-3,5), 8.02 (2H, d, J = 8.9Hz, H-2,6); ^{13}C NMR (d_6-DMSO) δ 198.2, 148.5, 145.1, 128.4 (CH), 123.1 (CH).

(b) Ethyl 2-(4-nitrophenyl)thiazole-4-carboxylate (2)

A suspension of the 4-nitrothiobenzamide (**1**)(3.5g, 19.2mmol) was stirred in ethanol (50mL) and ethyl bromopyruvate (3.75g, 19.2mmol, 1.0equiv.) was added. The mixture was heated at reflux for 4 hours then cooled and triethylamine (2.67mL, 19.2mmol, 1.0equiv.) added. The precipitate was collected on a filter, washed with water and dried under vacuum. The yield of white solid **2** was 4.01g (75%).

^{1}H NMR (d_6-DMSO) δ 8.69 (1H, s, H-5), 8.31 (2H, d, J = 9.0Hz, H-3′,5′), 8.20 (2H, d, J = 9.0Hz, H-2′,6′), 4.34 (2H, q, J = 7.1Hz, CH_2), 1.33 (3H, t, J = 7.1Hz, CH_3); ^{13}C NMR (d_6-DMSO) δ 165.1, 160.4, 148.3, 147.5, 137.6, 131.1 (CH), 127.5 (CH), 124.5 (CH), 61.0 (CH_2), 14.1 (CH_3).

(c) Ethyl 2-(4-aminophenyl)thiazole-4-carboxylate (3)

A solution of ethyl 2-(4-nitrophenyl)thiazole-4-carboxylate (**2**)(3.8g, 13.7mmol) and ammonium formate (5.17g, 82mmol) in ethanol (200mL) was stirred at room temperature. To this was added a suspension of 10%w/w palladium on charcoal (1.14g, 30%w/w) in ethanol (50mL). The reaction mixture was stirred for 36 hours at room temperature then filtered through celite. The celite pad was washed with hot ethanol (2 x 50mL). The combined filterates were concentrated to give a cream coloured crystalline solid **3** (3.65g). This was washed with water (3 x 50mL) and dried under vacuum. The yield was 2.252g, 66%).

^{1}H NMR (d_6-DMSO) δ 8.32 (1H, s, H-5), 7.65 (2H, d, J = 8.6Hz, H-2′,6′), 6.65 (2H, d, J = 8.6Hz, H-3′,5′), 5.77 (2H, bs, N-H), 4.32 (2H, t, J = 7.2Hz, CH_2), 1.33 (3H, t, J = 7.2Hz, CH_3); ^{13}C NMR (d_6-DMSO) δ 168.9, 160.9, 151.5, 146.4, 127.8 (CH), 126.4 (CH), 119.9, 113.6 (CH), 60.6 (CH_2), 14.2 (CH_3).

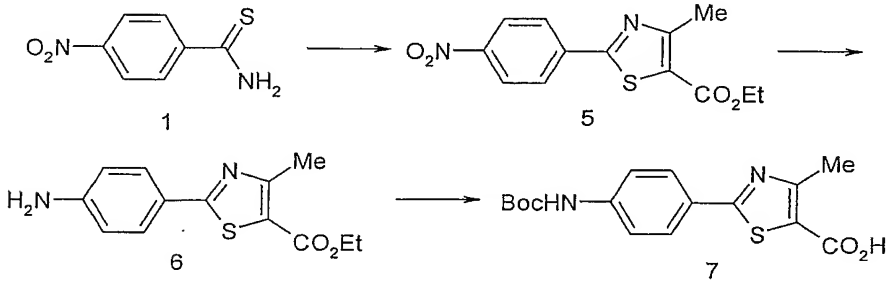
(d) 2-(4-tert-Butyloxycarbonylaminophenyl)thiazole-4-carboxylic acid (4)

Ethyl 2-(4-aminophenyl)thiazole-4-carboxylate (**3**)(1g, 4.0mmol) was dissolved in dry THF (25mL) and Boc anhydride (0.88g, 4.0mmol,

1.0equiv.) was added. The reaction mixture was heated at reflux under a nitrogen atmosphere for 18 hours. A further equivalent of Boc anhydride (0.88g, 4.0mmol) was then added and the mixture heated for a further 18 hours. The reaction mixture was cooled to room temperature and the solvent removed under vacuum. The residue was diluted with methanol (50mL) and then 1M aqueous sodium hydroxide solution (50mL) was added. The reaction mixture was heated at reflux for 4 hours then cooled to room temperature and stirred overnight. The volume was reduced under vacuum and the aqueous solution acidified with 1M hydrochloric acid (~50mL) to pH2-3. The resulting aqueous suspension was extracted with dichloromethane (4 x 50mL). The combined organic layers were dried over magnesium sulphate then concentrated under vacuum to give a pale yellow solid **4**, 1.27g (98%).

^{1}H NMR (d_6-DMSO) δ 13.10 (1H, bs, O-H), 9.70 (1H, bs, N-H), 8.42 (1H, s, H-5), 7.88 (2H, d, J = 8.7Hz, H-2',6'), 7.62 (2H, d, J = 8.7Hz, H-3',5'), 1.50 (9H, s, t-Bu CH_3); ^{13}C NMR (d_6-DMSO) δ 167.3, 162.0, 152.5, 147.9, 141.9, 127.9 (CH), 127.1 (CH), 126.3, 118.2 (CH), 79.5, 28.0 (CH_3).

Example 2 - 2-(4-tert-Butyloxycarbonylaminophenyl)thiazole-4-methyl-5-carboxylic acid (7)



(a) Ethyl 2-(4-nitrophenyl)thiazole-4-methyl-5-carboxylate (5)

This was made from 4-nitrothiobenzamide (**1**) by reacting with ethyl 2-chloroacetoacetate in a similar way to Example 1(b).

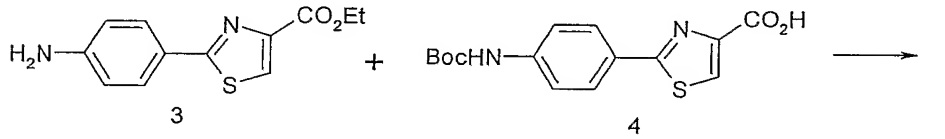
(b) Ethyl 2-(4-aminophenyl)thiazole-4-methyl-5-carboxylate (6)

This was made from ethyl 2-(4-nitrophenyl)thiazole-4-methyl-5-carboxylate (5) using the method of Example 1(c).

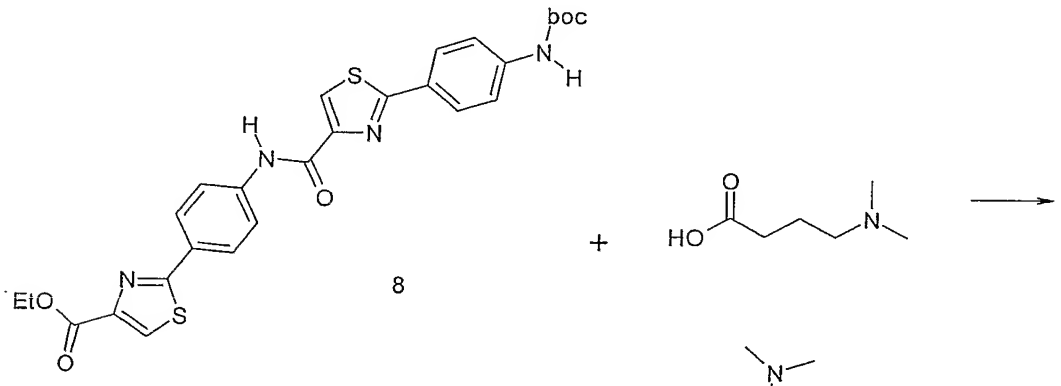
(c) 2-(4-tert-Butyloxycarbonylaminophenyl)thiazole-4-methyl-5-carboxylic acid (7)

This was made from ethyl 2-(4-aminophenyl)thiazole-4-methyl-5-carboxylate (6) using the method of Example 1(d).

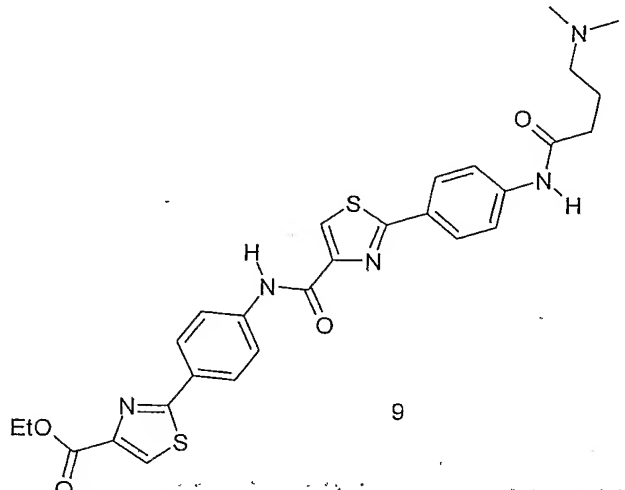
Example 3 - Ethyl 2-[4-({2-[4-(4-dimethylaminobutyrylamino)phenyl]thiazole-4-carbonyl}amino)phenyl]thiazole-4-carboxylate (9)



3 4



8



9

(a) Ethyl 2-(4-{[2-(4-tert-butyloxycarbonylaminophenyl)thiazole-4-carbonyl]amino}phenyl)thiazole-4-carboxylate (8)

Ethyl 2-(4-aminophenyl)thiazole-4-carboxylate (**3**)(0.039g, 0.16mmol) and 2-(4-tert-butyloxycarbonylaminophenyl)thiazole-4-carboxylic acid (**4**)(0.050g, 0.16mmol) were dissolved in dry DMF (1mL) and stirred under a nitrogen atmosphere. EDCI (0.060g, 0.16mmol, 2.0equiv.) and then DMAP (0.047g, 0.16mmol, 2.5equiv.) were added and the reaction mixture stirred at room temperature for 48 hours. The solution was diluted with ethyl acetate (10mL) and washed with 10%v/v hydrochloric acid (3 x 5mL) and then saturated sodium hydrogen carbonate solution (3 x 5mL). The organic layer was dried over magnesium sulphate then concentrated under vacuum to give an off white solid **8**. The yield was 0.070g (81%).

^{1}H NMR (d_6-DMSO) δ 10.44 (1H, bs, N-H), 9.73 (1H, bs, Boc N-H), 8.56 (1H, s, H-5), 8.46 (1H, s, H-5), 8.09 (2H, d, J = 8.8Hz, phenyl-H), 8.08 (2H, d, J = 8.8Hz, phenyl-H), 8.02 (2H, d, J = 8.8Hz, phenyl-H), 7.66 (2H, d, J = 8.8Hz, phenyl-H), 4.36 (2H, q, J = 7.1Hz, CH_2), 1.52 (9H, s, Boc CH_3), 1.35 (3H, t, J = 7.1Hz, CH_3).

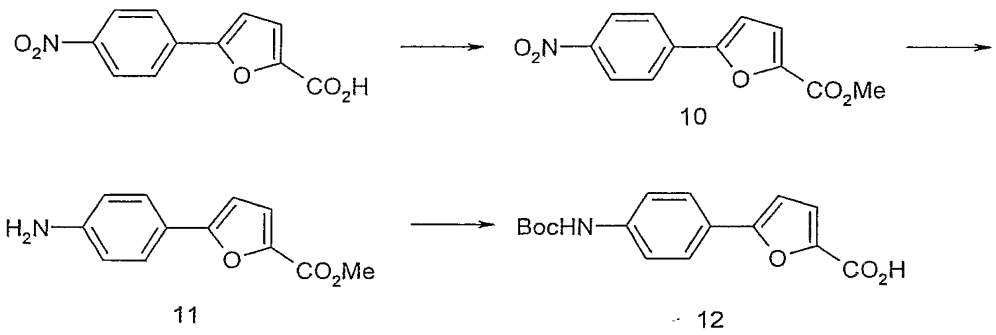
(b) Ethyl 2-[4-({2-[4-(4-dimethylaminobutyrylamino)phenyl]thiazole-4-carbonyl}amino)phenyl]thiazole-4-carboxylate (9)

Ethyl 2-(4-{[2-(4-tert-butyloxycarbonylaminophenyl)thiazole-4-carbonyl]amino}phenyl)thiazole-4-carboxylate (**8**)(0.020g, 0.036mmol) was dissolved in a 4M solution of hydrogen chloride in dioxane (1mL) with stirring. The reaction mixture was stirred for 1 hour under nitrogen, during which time a suspension formed. The solvent was removed under vacuum and the residue dried under vacuum. The residue and N,N-dimethylaminobutyric acid (0.016g, 0.12mmol, 3.3equiv) were dissolved in dry DMF (1mL) and stirred under a nitrogen atmosphere. EDCI (0.060g, 0.16mmol, 2.0equiv.) and then DMAP (0.047g, 0.16mmol, 2.5equiv.) were added and the reaction mixture stirred at room temperature for 96 hours. The solution was diluted with ethyl acetate (15mL) and washed with

saturated sodium hydrogen carbonate solution (3 x 5mL). The organic layer was dried over magnesium sulphate then concentrated under vacuum to give an off white solid **9**. The yield was 0.021g (93%).

^{1}H NMR (d_6-DMSO) δ 10.48 (1H, bs, N-H), 10.18 (1H, bs, N-H), 8.56 (2H, s, H-5), 8.20-7.98 (8H, m, phenyl-H), 4.36 (2H, t, J = 7.1Hz, CH_2), 2.45-1.53 (6H, m, butyryl C-H), 2.17 (3H, s, N-CH_3), 2.12 (3H, s, N-CH_3), 1.36 (3H, t, J = 7.1Hz, CH_3).

Example 4 - 5-(4-tert-Butoxycarbonylaminophenyl)-furan-2-carboxylic acid (12)



(a) Methyl 5-(4-nitrophenyl)-2-furoate (10)

A suspension of 5-(4-nitrophenyl)-2-furoic acid (4.9g, 21.0mmol) was suspended in dry dichloromethane (50mL) and oxalyl chloride (2.998g, 23.6mmol, 1.1equiv.) was added with stirring. After 5 minutes DMF (2 drops) was added and the flask fitted with a calcium chloride drying tube. The reaction mixture was stirred overnight during which time a homogeneous solution formed. A solution of triethylamine (4.768g, 46.2mmol, 2.2 equiv.) in dry methanol (20mL) was added dropwise to the acid chloride over 30 minutes. The reaction mixture was stirred for a further two hours then the concentrated under vacuum. The residue was taken up in ethyl acetate (200mL) and washed with 1M hydrochloric acid (3 x 50mL) and saturated sodium hydrogen carbonate solution (3 x 50mL). The organic layer was dried over magnesium sulphate then

concentrated under vacuum to a cream-coloured solid **10**, 4.972g (96%).

^{1}H NMR (d_6-DMSO) δ 8.32 (2H, d, J = 9.0Hz, H-3',5'), 8.06 (2H, d, J = 9.0Hz, H-2',6'), 7.48 (2H, s, H-3,4), 3.88 (3H, s, OCH_3); ^{13}C NMR (d_6-DMSO) δ 158.1, 154.2, 147.0, 144.5, 135.2, 134.5, 125.3 (CH), 124.5 (CH), 120.5 (CH), 111.6 (CH), 52.0 (CH_3).

(b) Methyl 5-(4-aminophenyl)-2-furoate (11)

A solution/suspension of methyl 5-(4-nitrophenyl)-2-furoate (**10**)(5.083g, 20.6mmol) in ethyl acetate (240mL) was added a suspension of 10% palladium on charcoal (0.5g, 10%equiv.) in ethyl acetate (10mL). The mixture was agitated under a hydrogen atmosphere (30psi) for 4 hours, then filtered through a celite pad. The celite was washed with ethyl acetate (2 x 50mL) and the combined filtrates concentrated under vacuum to give a pale yellow solid **11**, 3.905g (87%).

^{1}H NMR (d_6-DMSO) δ 7.50 (2H, d, J = 8.6Hz, H-2',6'), 7.34 (1H, d, J = 3.4Hz, H-3), 6.79 (1H, d, J = 3.6Hz, H-4), 6.66 (2H, d, J = 8.6Hz, H-3',5'), 5.59 (2H, bs, N-H), 3.82 (3H, s, OCH_3); ^{13}C NMR (d_6-DMSO) δ 158.7, 158.4, 150.0, 141.1, 125.9 (CH), 120.9 (CH), 116.5, 113.7 (CH), 104.0 (CH), 51.5 (CH_3).

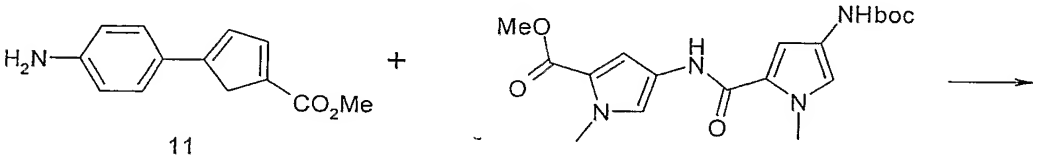
(c) 5-(4-tert-Butoxycarbonylaminophenyl)-furan-2-carboxylic acid (12)

To the methyl 5-(4-nitrophenyl)-2-furoate (**11**)(1.5g, 6.1mmol) dissolved in ethyl acetate (100mL) was added a suspension of 10% palladium on charcoal (0.3g, 20%equiv.) in ethyl acetate (20mL). The mixture was shaken under a hydrogen atmosphere (20psi) for 4 hours, then filtered through a celite pad. The celite was washed with ethyl acetate (2 x 20mL) and the combined filterates concentrated under vacuum. The residue was dissolved in dry THF (40mL) and Boc anhydride (1.323g, 6.1mmol, 1.0equiv.) was added. The mixture was stirred at room temperature then heated at reflux for 6 hours. The solvent was removed under vacuum and the residue dissolved in ethyl acetate (100mL). The solution was washed with

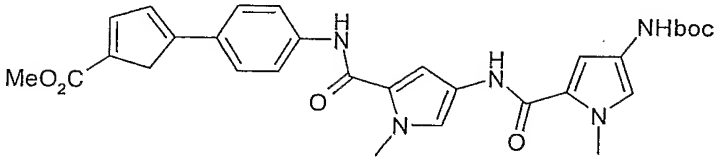
1M hydrochloric acid (3 x 50mL), then water (1 x 50mL). The organic layer was dried over magnesium sulphate then concentrated under vacuum to give a yellow solid. This was dissolved/suspended in methanol (50mL) and 1M sodium hydroxide (100mL) was added, then the mixture was heated at reflux for 4 hours then cooled and the methanol removed under vacuum. The solution was acidified with 1M hydrochloric acid (~100mL) then extracted with ethyl acetate (3 x 100mL). The combined organic extracts were dried over magnesium sulphate and then concentrated under vacuum to give a yellow solid **12**, 1.454g, (76%).

^{1}H NMR (d_6-DMSO) δ 9.57 (1H, bs, N-H), 7.69 (2H, d, J = 8.7Hz, H-2′,6′), 7.56 (2H, d, J = 8.7Hz, H-3′,5′), 7.28 (1H, d, J = 3.6Hz, H-3), 6.98 (1H, d, J = 3.6Hz, H-4), 1.48 (9H, s, Boc-CH_3); ^{13}C NMR (d_6-DMSO) δ 159.2, 156.5, 152.6, 143.4, 140.2, 125 (CH), 123.0, 120.0 (CH), 118.2 (CH), 106.4 (CH), 79.3, 28.0 (CH_3).

Example 5 - Methyl 4′-[(4-{[4-(4-butoxycarbonylamino)-1-methyl-1H-pyrrole-2-carbonyl]amino}-1-methyl-1H-pyrrole-2-carbonyl)amino]phenyl-2-furoate (13)



11



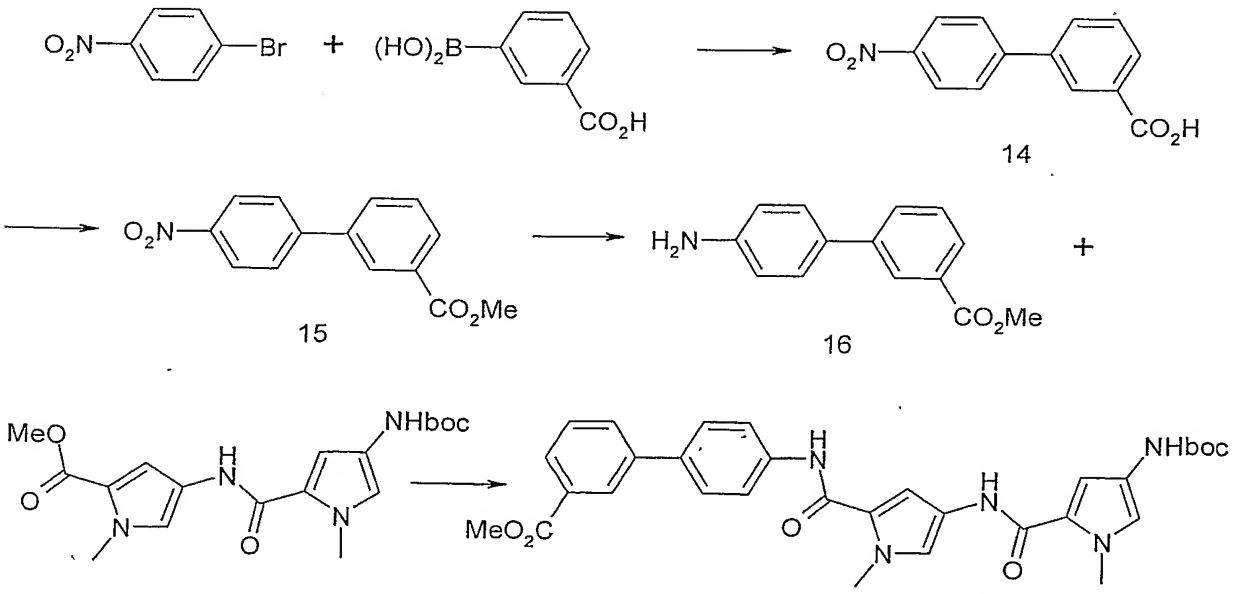
13

A solution of methyl 5-(4-aminophenyl)-2-furoate (**11**)(0.085g, 0.39mmol) and methyl 4-[(4-tert-butoxycarbonylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-1-methyl-1H-pyrrole-2-carboxylate (0.15g, 0.41mmol, 1.05equiv.) in dry DMF (2mL) was added a suspension of EDCI (0.159g, 0.83mmol, 2.0equiv.)in dry DMF (1mL) and dry dichloromethane (1mL) followed by DMAP (0.126g, 1.03mmol, 2.5equiv.) dissolved in dry DMF (0.5mL). The reaction mixture was

stirred at room temperature for 8 days then concentrated under vacuum to a volume of ~1mL. The residue was taken up in ethyl acetate (20mL) and washed with 10%v/v hydrochloric acid (3 x 10mL), then saturated sodium hydrogen carbonate solution (3 x 10mL). The organic fraction was dried over magnesium sulphate then concentrated under vacuum to give a yellow oil **13**, which was purified by column chromatography (silica gel, chloroform/methanol 1/99).

^{1}H NMR (d_6-DMSO) δ 10.05 (1H, bs, N-H), 9.90 (1H, bs, N-H), 9.10 (1H, bs, N-H), 7.90 (2H, d, J = 8.8Hz, H-2′,6′), 7.78 (2H, d, J = 8.8Hz, H-3′,5′), 7.43 (1H, d, J = 3.6Hz, py-H), 7.31 (1H, d, J = 1.5Hz, py-H), 7.21 (1H, d, J = 1.4Hz, py-H), 7.09 (1H, d, J = 3.6Hz, py-H), 6.92 (1H, bs, N-H), 6.87 (1H, bs, N-H), 3.88 (3H, s, OCH_3), 3.86 (3H, s, OCH_3), 3.84 (3H, s, OCH_3), 1.48 (9H, s, Boc-CH_3).

Example 6 – Methyl 4′-[(4-{[4-(4-butoxycarbonylamino)-1-methyl-1H-pyrrole-2-carbonyl]amino}-1-methyl-1H-pyrrole-2-carbonyl)amino]biphenyl-3-carboxylate (17)



(a) 3-(4-Nitrophenyl)benzoic acid (14)

1-Bromo-4-nitrobenzene (1.95g, 9.6mmol) and 3-carboxybenzeneboronic acid (1.8g, 10.8mmol, 1.1equiv.) were dissolved in a mixture of toluene (40mL), ethanol (40mL) and water (5mL) and potassium carbonate (4.1g, 29.3mmol, 3.0equiv.) was added. The flask was purged with nitrogen gas then palladium tetrakis(triphenylphosphine) (0.2g) was added and the mixture heated at reflux for 48 hours. The reaction mixture was cooled to room temperature and diluted with ethyl acetate (100mL) and extracted with water (3 x 50mL). The aqueous extracts were combined and washed with dichloromethane (3 x 50mL). The aqueous fraction was acidified (pH1-2) with concentrated hydrochloric acid to give an off white precipitate, which was collected on a filter and dried under vacuum to give a white solid **14**, 2.34g, (100%). ^{1}H NMR (d_6-DMSO) δ 13.20 (1H, bs, OH), 8.36-8.29 (3H, m, Ar-H), 8.07-8.01 (4H, m, Ar-H), 7.69 (1H, d, J = 7.8Hz, H-6); ^{13}C NMR (d_6-DMSO) δ 166.9, 146.9, 145.6, 138.2, 131.7 (CH), 131.6, 129.6 (CH), 128.0 (CH), 127.8 (CH), 124.1.

(b) Methyl 3-(4-nitrophenyl)benzoate (15)

A suspension of 3-(4-nitrophenyl)benzoic acid (**14**)(2g, 8.2mmol) in dry dichloromethane (50mL) and oxalyl chloride (1.15g, 9.0mmol, 1.1equiv.) was added. After 5 minutes DMF (2 drops) was added and the flask fitted with a calcium chloride drying tube. The reaction mixture was stirred overnight during which time a homogeneous solution formed. A solution of triethylamine (1.815g, 17.9mmol, 2.2 equiv.) in dry methanol (10mL) was added dropwise to the acid chloride over 20 minutes. The reaction mixture was stirred for a further two hours then the concentrated under vacuum. The residue was taken up in ethyl acetate (150mL) and washed with 1M hydrochloric acid (3 x 50mL) and saturated sodium hydrogen carbonate solution (3 x 50mL). The organic layer was dried over magnesium sulphate then concentrated under vacuum to a cream coloured solid **15**, 1.966g (88%).

^{1}H NMR (d_6-DMSO) δ 8.33 (2H, d, J = 8.9Hz, H-3′,5′), 8.27 (1H, m, Ar-H), 8.13-8.04 (2H, m, Ar-H), 8.01 (2H, d, J = 8.9Hz, H-2′,6′),

7.71 (1H, dd, J = 7.8Hz, H-5), 3.92 (3H, s, OCH_3); ^{13}C NMR (d_6-DMSO) δ 165.9, 147.0, 145.4, 138.3, 132.0 (CH), 130.6, 129.8 (CH), 129.5 (CH), 128.1 (CH), 127.6 (CH), 124.1 (CH), 52.3 (CH_3).

(c) Methyl 3-(4-aminophenyl)benzoate (16)

A solution of methyl 3-(4-aminophenyl)benzoate (**15**)(1.85g, 6.8mmol) in ethyl acetate (100mL) was added a suspension of 10% palladium on charcoal (0.185g, 10%equiv.) in ethyl acetate (10mL). The mixture was agitated under a hydrogen atmosphere (30psi) for 4 hours, then filtered through a celite pad. The celite was washed with ethyl acetate (3 x 50mL) and the combined filtrates concentrated under vacuum to give a pale yellow solid **16**, 1.663g (100%).

^1H NMR (d_6-DMSO) δ 8.09 (1H, dd, J = 1.7Hz, H-2), 7.85-7.80 (2H, m, H-4,6), 7.54 (1H, dd, J = 7.8Hz, H-5), 7.41 (2H, d, J = 8.5Hz, H-2',6'), 6.68 (2H, d, J = 8.5Hz, H-3',5'), 3.89 (3H, s, OCH_3); ^{13}C NMR (d_6-DMSO) δ 166.4, 148.9, 141.2, 130.1 (CH), 129.2 (CH), 127.3 (CH), 126.2 (CH), 126.0, 125.6 (CH), 114.3 (CH), 52.1 (CH_3).

(d) Methyl 4'-[(4-{[4-(4-butoxycarbonylamino)-1-methyl-1H-pyrrole-2-carbonyl]amino}-1-methyl-1H-pyrrole-2-carbonyl)amino]biphenyl-3-carboxylate (17)

A solution of methyl 3-(4-aminophenyl)benzoate (**16**)(0.095g, 0.39mmol) and methyl 4-[(4-tert-butoxycarbonylamino-1-methyl-1H-pyrrole-2-carbonyl)-amino]-1-methyl-1H-pyrrole-2-carboxylate (0.15g, 0.41mmol, 1.05equiv.) in dry DMF (2mL) was added a suspension of EDCI (0.159g, 0.83mmol, 2.0equiv.) in dry DMF (1mL) and dry dichloromethane (1mL) followed by DMAP (0.126g, 1.03mmol, 2.5equiv.) dissolved in dry DMF (0.5mL). The reaction mixture was stirred at room temperature for 8 days then concentrated under vacuum to a volume of ~1mL. The residue was taken up in ethyl acetate (20mL) and washed with 10%v/v hydrochloric acid (3 x 10mL), then saturated sodium hydrogen carbonate solution (3 x 10mL). The organic fraction was dried over magnesium sulphate then concentrated under vacuum to give a yellow oil **17**, which was

purified by column chromatography (silica gel, chloroform/methanol 1/99).

^{1}H NMR ($CDCl_3$) δ 8.28 (1H, dd, J = 1.6Hz, Ar-H), 8.00 (1H, dd, J = 1.3, 7.8Hz, Ar-H), 7.78 (1H, m, Ar-H), 7.69-7.60 (6H, m, Ar-H), 7.51 (1H, dd, J = 1.7Hz, H-5 (biphenyl)), 7.45 (1H, bs, N-H), 7.15 (1H, d, J = 1.7Hz, py-H), 6.82 (1H, d, J = 1.8Hz, py-H), 6.61 (1H, bs, N-H), 6.22 (1H, bs, N-H), 3.98 (3H, s, CH_3), 3.96 (3H, s, CH_3), 3.93 (3H, s, CH_3), 1.55 (9H, s, Boc-CH_3).

Example 7 – DNA Footprinting

In order to assess the binding of test compounds to DNA, a footprinting study against the MS2 DNA sequence was carried out. The sequence is as follows:

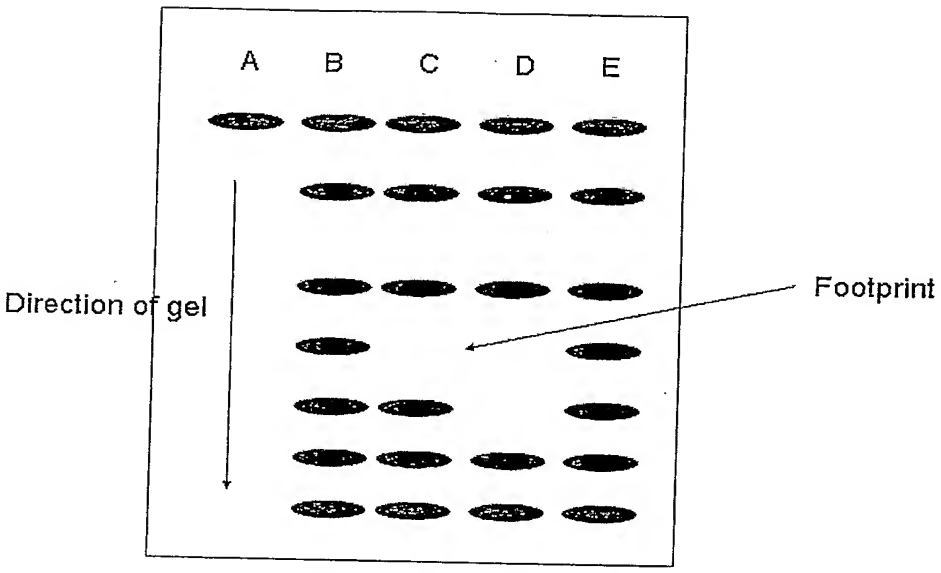
5′- CAGGAGGCAG CTATGACCAT GATTACGAAT TCGAGCTCGG TACCCGGGGA
TCCATATGCG GCAATACACA TGGCCGATTT CCAACGTCAC TAGTCGTAGC
GCGATCAAGG TTAAGCTCCC GTTCTATCCT GGTATAGCAA TTAGGGCGTG
AAGAGTTATG TAAAGTACGT CCGGTGGGGT CTGTTTTGTC ATCTCAGCCT
CGAATGCGGA TCCTCTAGAG TCGACCTGCA GGCATGCAAG CTTGGCACTG
GCCGTCGTTT TA -3′

and is derived from a bacteriophage.

The footprinting technique in the context of DNA allows the determination of binding sites for drugs or biological macromolecules. It relies on the fact that DNA is cleaved relatively non-specifically by free radicals (e.g. hydroxyl radical footprinting) or enzymes (e.g. DNase I footprinting). Thus if DNA is ^{32}P end labelled on one strand (so that it may be observed autoradiographically) and exposed to the cleavage agent for a certain time, then a laddering pattern may be seen when the resulting fragments are separated by gel electrophoresis (separation on the basis of size). If a compound which binds to DNA is added prior to the cleavage agent then this hinders access of the enzyme or radical to the DNA and blocks cleavage at the

molecule binding site. Thus if the compound binds discretely then a specific cleavage block should be seen on the gel relative to the DNA not treated with compound, which is termed the 'footprint'. (see Figure 1)

Compound 9 produces an unusual profile in which there appears to be no specific footprinting activity, but conversely there is no clear coating event. There may be a structural aspect to the cleavage pattern seen.



A – DNA – no enzyme (no cleavage pattern), B – DNA + enzyme (control), C – DNA + compound (low conc.) + enzyme, D – DNA + compound (high conc.) + enzyme, E – DNA + enzyme (control).

Fig. 1